| | |
|---|---|
| **Titre:** Title: | Convolution et déconvolution linéaire 3-D appliquées à l'holographie acoustique temporelle |
| **Auteur:** Author: | Jean-Michel Attendu |
| **Date:** | 2017 |
| **Type:** | Mémoire ou thèse / Dissertation or Thesis |
| **Référence:** Citation: | Attendu, J.-M. (2017). Convolution et déconvolution linéaire 3-D appliquées à l'holographie acoustique temporelle [Thèse de doctorat, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/2932/ |



| | |
|---|---|
| **URL de PolyPublie:** PolyPublie URL: | https://publications.polymtl.ca/2932/ |
| **Directeurs de recherche:** Advisors: | Annie Ross |
| **Programme:** Program: | Génie mécanique |

UNIVERSITÉ DE MONTRÉAL

# CONVOLUTION ET DÉCONVOLUTION LINÉAIRE 3-D APPLIQUÉES À L'HOLOGRAPHIE ACOUSTIQUE TEMPORELLE

JEAN-MICHEL ATTENDU

DÉPARTEMENT DE GÉNIE MÉCANIQUE

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION

DU DIPLÔME DE PHILOSOPHIAE DOCTOR

(GÉNIE MÉCANIQUE)

DÉCEMBRE 2017

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée :

# CONVOLUTION ET DÉCONVOLUTION LINÉAIRE 3-D APPLIQUÉES À L'HOLOGRAPHIE ACOUSTIQUE TEMPORELLE

présentée par : ATTENDU Jean-Michel

en vue de l'obtention du diplôme de : Philosophiae Doctor

a été dûment acceptée par le jury d'examen constitué de :

M. KHAMENEIFAR Farbod, Ph. D., président

Mme ROSS Annie, Ph. D., membre et directeur de recherche

M. CARDINAL Christian, Ph. D., membre

M. BERRY Alain, Ph. D., membre externe

# DÉDICACE

*En tant que fils, petit-fils et arrière-petit-fils d'ingénieurs, je dédie cette thèse à mes prédécesseurs Claude, Michel et André-Claude qui ont su transmettre leur curiosité scientifique au-delà des générations.*

# REMERCIEMENTS

Mes premiers remerciements vont à ma directrice de recherche, la professeure Annie Ross, pour sa précieuse confiance dont j'ai bénéficié tout au long du projet de recherche. Mme Ross a toujours été à l'écoute de mes interrogations et de mes propositions et a été très généreuse de son temps. Elle a su me proposer plusieurs opportunités qui m'ont permis de m'épanouir à travers mon projet doctoral. Son mentorat et ses conseils ont contribué de manière significative aux apprentissages que j'ai faits au cours des dernières années et ainsi qu'à la réussite de cette thèse. Tant sur l'aspect technique que sur l'aspect humain, Annie Ross restera pour moi un modèle dont l'influence laissera sans doute une emprunte durable.

Je tiens également à remercier ma famille et mes amis pour leur soutien indéfectible. Parents, frère, sœurs, oncles et tantes, grands-parents, cousins m'ont tous, d'une manière ou d'une autre, démontré leur soutient et leurs encouragements au cours de mon doctorat. Ma fierté et ma persévérance découlent en bonne partie de ces encouragements. En ce sens, je remercie particulièrement Valérie, avec qui je partage mon quotidien et qui est ma principale confidente.

Je remercie également Jean-Hugh Thomas pour m'avoir chaleureusement accueilli dans son laboratoire de juillet à décembre 2015 et les doctorants du LAUM pour leur vive camaraderie : je garde un excellent souvenir de mon séjour en France. Les connaissances partagées, notamment sur les problèmes inverses, ont été fondamentales pour le développement de certaines de mes idées.

Je remercie également mes collègues et amis de Polytechnique pour les mille et une activités que nous avons partagées au cours des dernières années. Mes remerciements vont particulièrement à mes collègues du LAVA, à ceux de mécanique appliquée, ainsi qu'à ceux de l'AECSP. Merci également au personnel du département pour leur aide et nos échanges cordiaux. Je remercie aussi mes stagiaires qui m'ont donné une aide appréciée avec la partie expérimentale du projet.

Finalement, je remercie les réviseurs de mes articles et mon jury de thèse pour avoir accepté de réviser mes travaux. Je suis conscient que la compréhension de certains aspects plus mathématiques nécessite quelques efforts : votre expertise et votre investissement sont grandement appréciés.

# RÉSUMÉ

L'holographie acoustique en champ proche est une méthode permettant de calculer l'amplitude d'un champ sonore dans l'espace tridimensionnel et dans le temps. Son principe repose sur l'utilisation d'une antenne microphonique placée dans le champ proche de la source, qui permet d'obtenir une condition frontière suffisante à la résolution de l'équation d'onde avec le formalisme de Green. La formulation standard permet de déterminer le champ de sources stationnaires. Il est alors supposé que la source rayonne de manière périodique dans le temps : le champ reconstruit est conséquemment représenté dans le domaine fréquentiel. Lorsque la source est non-stationnaire, le champ doit être représenté dans le domaine temporel et le formalisme standard ne s'applique pas, car la nature finie et discrète du champ mesuré introduit plusieurs erreurs de traitement du signal, notamment à cause de l'utilisation de la transformée de Fourier discrète qui suppose la périodicité. Les principales erreurs engendrées sont dues aux fuites spectrales (*leakage*) et au phénomène de recouvrement (*wrap-around*). Dans cette thèse, nous proposons un formalisme permettant de supprimer ces erreurs en utilisant un opérateur de convolution linéaire tridimensionnel de même que l'utilisation d'une fonction de Green échantillonnée dans le domaine temps-espace. Cette formulation améliore significativement la reconstruction du champ sonore tant pour le problème direct (propagation) que le problème inverse (rétro-propagation).

En propagation, l'erreur obtenue avec la méthode proposée est environ dix fois moindre qu'avec les méthodes compétitrices dans les cas étudiés. De plus, on peut contrôler le niveau de l'erreur en choisissant un fenêtrage et une discrétisation adéquate. On établit également un lien entre l'erreur de troncature de la fonction de Green et la distance entre le plan de mesure et le plan de propagation. On valide la méthode développée en l'appliquant à des signaux simulés à partir d'un modèle de piston bafflé, ainsi qu'à des signaux expérimentaux obtenus dans un environnement anéchoïque.

En rétro-propagation, on obtient une réduction de l'erreur allant jusqu'à un facteur trois en comparaison à la méthode standard basée sur l'utilisation de la convolution circulaire. L'erreur reste significative puisque le problème inverse est mathématiquement mal posé. Par conséquent, on analyse la performance de la méthode de Tikhonov pour régulariser le problème selon le niveau de bruit ajouté et la distance de propagation. Trois différentes méthodes sont utilisées pour déterminer le paramètre de régularisation optimal, soient la méthode de la courbe en L, la validation croisée généralisée et une approche Bayésienne. La méthode de validation croisée généralisée est,

dans les cas étudiés, celle qui permet d'obtenir la meilleure prédiction du paramètre de régularisation optimal.

On analyse également les performances de la méthode de régularisation parcimonieuse, lorsqu'appliquée au formalisme d'holographie acoustique proposé. Cette technique est intéressante dans le cas de sources causales puisque l'amplitude du champ rayonné par ces sources est nulle avant que le front d'onde atteigne le point d'observation. Puisque la méthode de régularisation parcimonieuse favorise la convergence vers zéro de la solution, cette méthode permet de mieux reconstruire les sources causales dans ce régime. On valide cette hypothèse en appliquant cette méthode à des signaux simulés et expérimentaux. On démontre également que la méthode de la courbe de Pareto, une adaptation de la courbe en L, permet de bien prédire le paramètre de régularisation pour les cas étudiés.

Finalement, dans un contexte plus général, nous proposons des algorithmes basés sur l'utilisation de la transformée de Fourier généralisée permettant d'améliorer l'efficacité de calcul de la convolution linéaire. On propose une formule permettant de déterminer le coefficient de modulation optimal de la transformée de Laplace numérique pour le calcul de la convolution linéaire. On propose également une procédure qui permet de réduire les coûts de calcul de la convolution linéaire multidimensionnelle d'environ 20 % en dédoublant l'utilisation des coefficients de modulation avec la transformée de Fourier généralisée.

# ABSTRACT

Near-field acoustic holography is a method used to determine the amplitude of a sound field in three-dimensional space and in time. It is based on the use of a microphone array placed in the near field of the source, which is used to obtain a boundary condition sufficient to the resolution of the wave equation with Green's formalism. The standard formulation is used to determine the field of stationary sources. It assumes that the source radiates periodically in time: the reconstructed field is consequently represented in the frequency domain. When the source is non-stationary, the field must be represented in the time domain and the standard formulation does not apply, because the finite and discrete nature of the measured field introduces several signal processing errors, notably because of the use of the discrete Fourier transform which assumes the periodicity. The main errors generated are due to leakage and wrap-around phenomena. In this thesis, we propose a formalism to eliminate these errors by using a three-dimensional linear convolution operator as well as the use of a Green's function sampled in the time and spatial domain. This formulation significantly improves the reconstruction of the sound field both for the direct problem (propagation) and the inverse problem (retro-propagation).

In propagation, the error obtained with the proposed method is about ten times smaller than with the competing methods in the cases studied. In addition, the error can be controlled by choosing an adequate window and a proper sampling frequency. It is shown that the truncation error due to the Green's function is proportional to the distance between the measurement plane and the propagation plane. The proposed method is validated by applying it to simulated signals from a baffled piston model, as well as to experimental signals obtained in an anechoic environment.

In back propagation, the error is reduced by up to a factor three in comparison to the circular convolution based method. The error remains significant since the inverse problem is mathematically ill-posed. The performance of Tikhonov's method is analyzed to regularize the problem with respect to the noise level and the propagation distance. Three different methods are used to determine the optimal regulation parameter, namely the L-curve, the generalized cross-validation and a Bayesian approach. The generalized cross-validation method gives the best prediction of the optimal regulation parameter in the cases studied.

The performance of the sparse regularization method is also analyzed when applied to the proposed NAH formulation. This technique is interesting in the case of causal sources, since the amplitude

of the field radiated by these sources is null before the wave front reaches the point of observation. Since the sparse regularization method favors the convergence of the solution towards zero, it is assumed that this method produces an improved reconstruction of the causal source. This assumption is validated by applying this method to simulated and experimental signals. It is also shown that the Pareto frontier curve, an adaptation of the L curve, gives adequate prediction of the regularization parameter in the cases studied.

Finally, in a more general context, we propose algorithms based on the use of the generalized discrete Fourier transform to improve the computational efficiency of the linear convolution. A formula is proposed for determining the optimal modulation coefficient of the numerical Laplace transform for calculating linear convolution. An algorithm is also proposed to reduce the computational costs of the multidimensional linear convolution by about 20% by duplicating the use of the modulation coefficients with the generalized discrete Fourier transform.

# TABLE DES MATIÈRES

# LISTE DES TABLEAUX

# LISTE DES FIGURES

# LISTE DES ACCRONYMES

| | |
|---|---|
| DFT | Discrete Fourier transform |
| FFT | Fast Fourier transform |
| GCV | Generalized cross-validation |
| GDFT | Generalized discrete Fourier transform |
| HELS | Helmoltz equation least-square method |
| NAH | Nearfield acoustical holography |
| NLT | Numerical Laplace transform |
| RMS | Root mean square |
| RT-NAH | Real-time nearfield acoustical holography |
| TD-NAH | Time domain nearfield acoustical holography |
| TD-PWSM | Time domain plane wave superposition method |
| TDH | Time domain holography |
| TF | Transformée de Fourier |
| TNAH | Transient nearfield acoustical holography |
| TSD | Time-space domain |

# LISTE DES ANNEXES

# CHAPITRE 1 INTRODUCTION ET PLAN DE LA THÈSE

## 1.1 Introduction

Les signaux acoustiques non-stationnaires peuvent prendre plusieurs formes : bruits d'impact, coups de moteur, bruits de crissement, sons musicaux, bruits induits par les machines, signaux vocaux, etc. Leur abondance dans les environnements industriels, le transport, la construction et leur impact sur la santé et la sécurité sont autant de justifications de la pertinence d'avoir des outils permettant de visualiser ces champs sonores. L'holographie acoustique en champ proche (*Nearfield Acoustical Holography* – NAH) est une technique d'imagerie du champ sonore introduite au début des années 80 par Maynard et Williams [1]. Cette technique permet de déterminer la pression ou la vélocité particulaire dans un espace tridimensionnel à partir de mesures réalisées sur un espace 2-D, d'où son appellation d' « holographie ». Une particularité du NAH est que la méthode tire avantage de la contribution des ondes évanescentes au champ proche en effectuant la prise de mesure à proximité de la source. Dans le domaine de Fourier, les ondes évanescentes correspondent aux composantes à grands nombres d'ondes. Celles-ci permettent d'obtenir une meilleure résolution spatiale puisque ces composantes correspondent à de petites longueurs d'onde dans le domaine spatial [2, 3].

Cette prise de mesure dans le champ proche s'effectue avec une matrice de microphones. Cette mesure consiste en une condition frontière pour la résolution de l'équation d'onde avec le formalisme de Green [4]. On obtient alors une équation décrivant le champ sonore en quatre dimensions (fréquence-espace) selon la pression obtenue sur la surface. Celle-ci prend la forme d'un produit de convolution entre cette condition frontière et une fonction de transfert obtenue avec la méthode de Green [3]. Ce produit de convolution s'effectue selon trois dimensions et, dans le cas du NAH standard, est résolu dans l'espace de Fourier avec la propriété de convolution de la transformée de Fourier discrète (*Discrete Fourier Transform* - DFT) [2, 5]. Un des principaux défis en holographie en champ proche est de déterminer le champ sonore, malgré les erreurs de traitement du signal introduites par la nature discrète et finie du problème.

En effet, parce que la condition frontière consiste en une quantité finie de données, cette technique comporte certaines limitations lorsqu'appliquée à l'étude de signaux non périodiques : l'utilisation de la DFT étant basée sur l'hypothèse de périodicité du signal. Il en résulte la génération d'erreurs

de recouvrement (« *wrap-around* ») qui faussent les résultats, ainsi que des problèmes de fuites spectrales (« *leakage* » ). Plusieurs reformulations du NAH ont été développées de manière à réduire ces erreurs. Cependant, ces techniques ont des marges d'erreur considérables, surtout dans le cas où la distance entre la mesure et la reconstruction est grande [6, 7, 8].

Un autre problème majeur en NAH est la régularisation du problème inverse, c'est-à-dire le traitement des instabilités mathématiques qui surviennent lors de la reconstruction du champ sonore entre la mesure et la source. En effet, le champ mesuré est réamplifié puisque l'onde s'atténue en s'éloignant de la source. Cette amplification est particulièrement importante dans le cas des composantes évanescentes dont l'amplitude décroit exponentiellement en s'éloignant de la source. Le problème devient alors très sensible à l'introduction de bruit de mesure : on dit alors que le problème est mal posé [9, 10]. Il convient alors d'appliquer des techniques de régularisation pour limiter l'amplification du bruit. La régularisation est un problème ouvert en mathématique : l'existence (ou l'inexistence) d'une technique de régularisation optimale n'est pas démontrée. Plusieurs méthodes existent, mais leurs performances dépendent de la fonction de Green, de la mesure et du bruit ajouté [9]. Ainsi, pour chaque formulation du NAH, il convient de tester les différentes méthodes de régularisation afin de déterminer laquelle performe le mieux et selon quelles circonstances.

Dans le cadre de ce projet doctoral, on propose le développement d'une variante du NAH, la formulation temps-espace, qui supprime entièrement les erreurs de *wrap-around* et de fuites spectrales lorsqu'appliquée aux signaux non-stationnaires. On suppose que les sources non-stationnaires étudiées sont causales et que leur amplitude devient nulle au-delà d'un temps $t$ fini. L'élément clé de la formulation est l'utilisation de la convolution linéaire 3-D, qui suppose des valeurs nulles à l'extérieur du champ mesuré. La méthode proposée permet aussi de conserver une grande efficacité de calcul puisqu'elle tire profit de l'algorithme de transformée de Fourier rapide (*Fast Fourier transform* – FFT). Une telle formulation permet de visualiser les signaux non-stationnaires avec une précision supérieure aux méthodes actuellement utilisées. De plus, l'application des méthodes de régularisation $\ell^2$ et $\ell^1$ à la résolution du problème inverse avec la formulation proposée est analysée. Finalement, deux algorithmes de calcul de la convolution linéaire permettant une plus grande efficacité de calcul sont présentés.

## 1.2 Plan de la thèse

Cette thèse suit le modèle de thèse par article. Les premiers chapitres servent donc à mettre en contexte l'étude et à fournir au lecteur certaines notions permettant de l'aborder avec les connaissances nécessaires. Ainsi, suite à cette introduction, un chapitre décrivant des notions théoriques préliminaires est présenté et suivi d'une description de l'état de l'art en holographie acoustique temporelle. Une première partie des notions préliminaires est présentée en annexe, étant donnée leur nature plus fondamentale. L'état de l'art permet de définir les objectifs de la thèse au chapitre suivant. La méthodologie employée pour l'atteinte de ces objectifs est également décrite. Ensuite, les trois articles formant le cœur de la thèse sont présentés. Le premier article présente le formalisme mathématique et analyse les performances de la méthode temps-espace lorsqu'appliqué au problème direct. Le deuxième article porte sur l'application du formalisme temps-espace au problème inverse, l'utilisation d'un algorithme permettant de supprimer certaines erreurs de troncature, l'application de la régularisation de Tikhonov ainsi que sur trois formules permettant de prédire le paramètre de régularisation. Finalement, le troisième article étudie l'application de la régularisation parcimonieuse à la formulation temps-espace et démontre l'efficacité de la courbe de Pareto pour prédire le paramètre de régularisation. Dans l'ensemble, les trois articles présentent un développement complet du formalisme d'holographie acoustique temps-espace. La formulation proposée permet une détermination précise du champ acoustique non-stationnaire et ce, en propagation comme en rétro-propagation et pour divers types de signaux. Finalement, un chapitre de discussion effectue la synthèse des travaux présentés. Celui-ci est suivi d'un dernier chapitre exposant les principales conclusions et recommandations. Deux articles en marge des principaux objectifs de thèse sont présentés en annexe : ceux-ci traitent d'algorithmes permettant le calcul de la convolution linéaire de manière plus efficace. Un schéma des différentes sections de la thèse et des liens qui les unissent est présenté à la Figure 1.1.

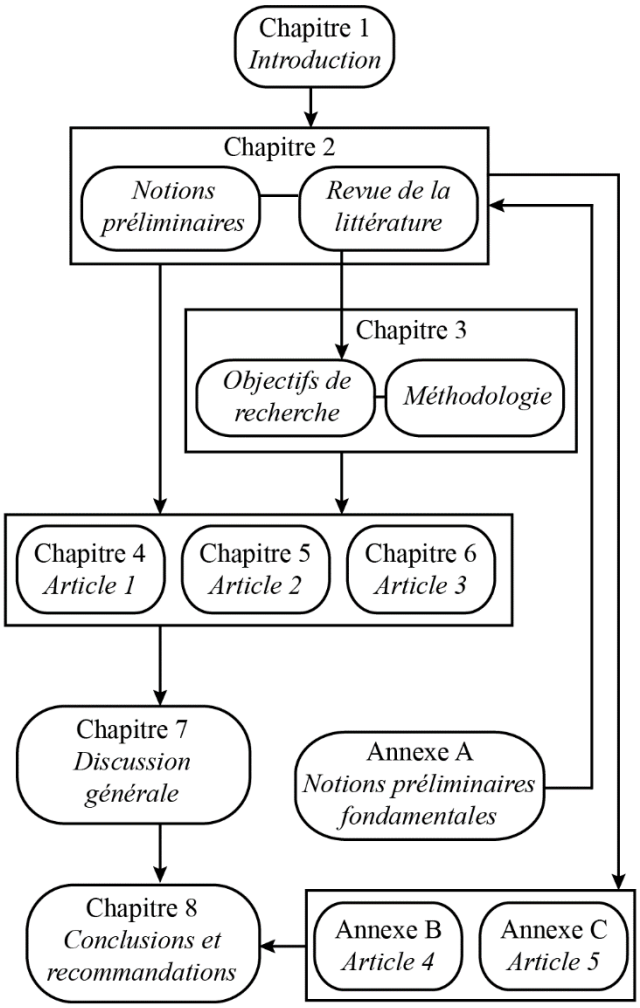


*Figure 1.1 : Plan de la thèse*

# CHAPITRE 2 REVUE DE LITTÉRATURE

Ce chapitre met en place les notions préliminaires permettant au lecteur d'aborder les articles avec les outils mathématiques et physiques nécessaires. Afin d'en alléger la lecture, une première partie des notions préliminaires, qui traite de sujets plus fondamentaux, est présentée à l'annexe A. À travers ces notions préliminaires, on développe le formalisme analytique derrière la méthode d'holographie acoustique temporelle. Ceci permet d'exposer les principales hypothèses inhérentes à la méthode. Ce chapitre aborde également les différents problèmes liés à la nature discrète et finie des signaux mesurés et introduit ainsi d'importantes problématiques résolues par ces travaux de thèse. Une revue de la littérature des différentes adaptations de l'holographie acoustique en champ proche appliquée à la détermination de champs rayonnés par des sources non-stationnaires est présentée. Une synthèse des principaux problèmes des différentes méthodes existantes est présentée. Finalement, ce chapitre permet également de définir plusieurs notations utilisées dans l'ensemble du document.

## 2.1 Formulation analytique de l'holographie acoustique en champ proche

On applique le formalisme de Green développé aux sections 9.4 à 9.6 de l'annexe A au problème de rayonnement acoustique. On suppose la géométrie planaire illustrée à la Figure 2.1, où $A$ est la source étudiée. On suppose que la source est bafflée, c'est-à-dire qu'elle ne rayonne que pour $z > 0$. On considère ainsi une surface fermée $S$ qui ne contient pas $A$, mais par laquelle passe la totalité du champ rayonné, ce qui permet l'application du théorème du gradient présenté à la section 9.3 de l'annexe A. La surface $S = D \cup H$ est délimitée par deux surfaces $D$ et $H$, où $D$ est un disque circulaire de rayon $R_0$ situé sur le plan $z = z_0$ parallèle au baffle et $H$ est la demi-sphère de rayon $R_0$ connectée au disque, comme le montre la Figure 2.1.

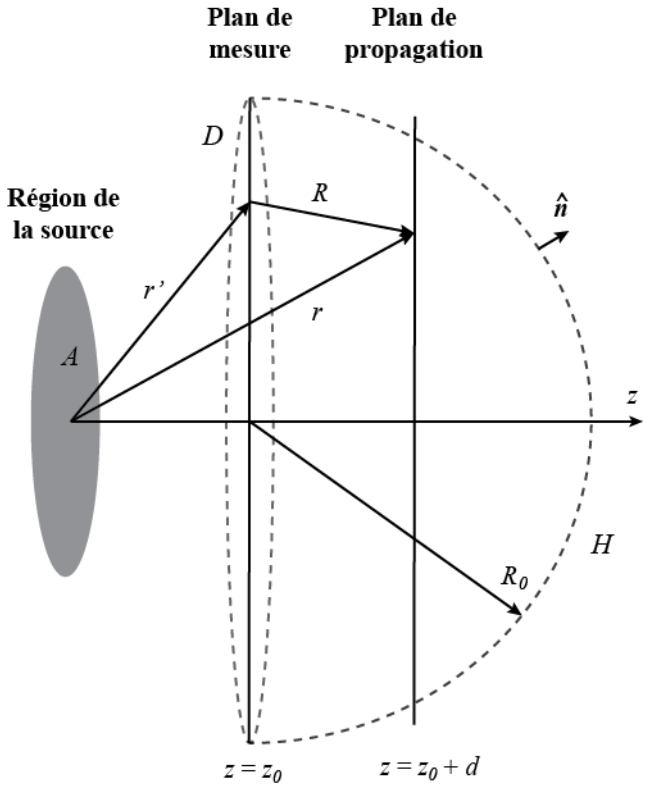


*Figure 2.1 : Géométrie considérée pour la détermination du champ de pression sur un plan $p(z > z_0)$ à partir du champ de pression au plan $p(z_0)$*

L'objectif de la méthode d'holographie acoustique est de mettre en relation la pression $p(r)$ au plan $z = z_0 + d$ dans le demi-espace $z > z_0$ avec le champ de pression $p(r')$ au plan $z = z_0$. En appliquant la deuxième identité de Green avec la surface fermée $S$, on obtient:

$$p(r) = \iint_S \left( p(r') \frac{\partial g_D}{\partial n} - g_D \frac{\partial p(r')}{\partial n} \right) dS. \tag{2.1}$$

Les variables $r$ et $r'$ de l'équation (A.29) ont été interchangées pour satisfaire à la notation présentée à la Figure 2.1. En prenant $R_0 \to \infty$, on retrouve une condition de Sommerfeld [11] et le résultat de l'intégrale surfacique sur $S$ s'annule. En effet, puisque le champ de pression diminue à mesure qu'il se propage en s'éloignant de la source, à une distance infinie de celle-ci, le champ est nul.

Sur la surface $D$, puisqu'on choisit la fonction de Dirichlet, on a que $g_D = 0$ et $\partial g_D / \partial n = 2\, \partial g / \partial z$, et l'équation (2.1) devient:

$$p(r) = 2 \iint_D p(r') \frac{\partial}{\partial z'} g(r, r')\, dx' dy'. \tag{2.2}$$

On considère ici une géométrie planaire : les surfaces de mesure et de propagation sont planes. Cependant, on peut adapter la résolution de l'équation (2.1) à d'autres types de géométrie [2]. Après avoir appliqué l'opérateur différentiel, on obtient :

$$p(r) = \iint_D p(r')(ik - 1/R) \frac{e^{ikR}}{R^2} dx' dy'. \tag{2.3}$$

On reconnait à l'équation (2.3) une intégrale de convolution 2-D sur $x$ et $y$, où $p(r')$ est le champ mesuré sur $D$, $p(r)$ est le champ déterminé au plan de propagation et la fonction de transfert $G$ est :

$$G(\omega, x, y, d) = \left(ik - \frac{1}{R}\right) e^{ikR}/R^2. \tag{2.4}$$

Étant donnée la propriété de convolution de la *TF*, la convolution sur le domaine temporel s'exprime comme un produit dans l'espace des fréquences $\omega$. En utilisant l'identité [1]:

$$2\pi \frac{\partial}{\partial z} \frac{e^{ikR}}{R} = - \iint_{-\infty}^{\infty} e^{i\left(k_x(x-x') + k_y(y-y') + k_z(z-z_0)\right)} dxdy, \tag{2.5}$$

On détermine l'expression de la fonction de transfert dans l'espace $(\omega, k_x, k_y)$ :

$$G\left(\omega, k_x, k_y, d\right) = e^{i \cdot d \cdot \sqrt{k^2 - {k_x}^2 - {k_y}^2}} = e^{i \cdot d \cdot k_z}. \tag{2.6}$$

L'équation (2.6) est la forme standard utilisée en holographie acoustique stationnaire [2, 1, 3].

Finalement, en considérant la propriété de différentiation et la propriété de glissement, présentées à la section 9.1 de l'annexe A, sur l'équation (2.4), on peut exprimer la fonction de transfert dans le domaine du temps-espace $(t, x, y)$ [12]:

$$G(t, x, y, d) = \frac{d}{2\pi c} \left(\frac{c}{R^3} + \frac{1}{R^2} \frac{\partial}{\partial t}\right) \delta(t - R/c). \tag{2.7}$$

Ainsi, le champ propagé peut être déterminé dans le domaine temps-espace (*time and spatial domain* – TSD) en effectuant la convolution 3-D entre le champ mesuré dans le TSD et la fonction de transfert :

$$p(t,x,y,z_0+d) = \iiint_{-\infty}^{\infty} p(t',x',y',z_0)\cdot G(t-t',x-x',y-y',d)dt\,dx'dy'. \qquad (2.8)$$

À partir des fonctions de Green exprimées aux équations (2.4), (2.6) et (2.7), on a trois formulations équivalentes au problème d'holographie acoustique. Lorsque le champ de pression est connu analytiquement sur tout le domaine, le champ propagé obtenu est le même avec toutes les approches. Cependant, lorsque le champ est obtenu sur un domaine fini et discret, des erreurs de traitement du signal sont introduites et les résultats des trois approches diffèrent.

## 2.2 Détermination du champ vectoriel

En utilisant l'équation d'Euler (A.15) qui relie la vitesse particulaire à la pression, on peut déterminer des fonctions de Green alternatives permettant d'obtenir l'intensité sonore, l'accélération ou la vitesse particulaire au plan de propagation [2, 13]. Typiquement en NAH, on intègre la transformation à même la fonction de transfert, puisque celle-ci est connue analytiquement. On peut faire cette transformation puisqu'un opérateur différentiel appliqué au résultat d'une convolution peut, de manière équivalente, être appliqué à l'une ou l'autre des fonctions convoluées puisque la convolution est commutative [14]. On obtient ainsi une fonction de transfert qui, convoluée avec le champ de pression mesuré, produit le champ vectoriel de vitesse particulaire au plan de propagation. La relation d'Euler permet d'obtenir le vecteur de vitesse particulaire et l'intensité sonore peut être obtenue à partir du produit de la pression et de la vitesse. L'accélération est obtenue en dérivant la vitesse selon $t$.

Dans le cas de la fonction de transfert standard $G(\omega, k_x, k_y)$ exprimée dans l'espace fréquence-nombres d'onde, les fonctions de transfert correspondantes sont présentées au tableau 2.1.

*Tableau 2.1: Fonctions de transfert pour obtenir différentes quantités physiques au plan de propagation*

| Quantités physiques | Fonctions de transfert |
|---|---|
| Pression | $G = e^{ik_z(z-z_0)}$ |
| Vitesse particulaire | $G_v = \frac{k_z}{\rho_0 ck} e^{ik_z(z-z_0)}$ |
| Accélération | $G_a = \frac{-ik_z}{\rho_0} e^{ik_z(z-z_0)}$ |
| Intensité | $G_I = \frac{1}{2} Re\{G \cdot \overline{G_v}\}$ |

Les équations présentées sont valides pour une source planaire. Des développements similaires sont proposés dans la littérature pour des sources de géométrie cylindrique, sphérique ou quelconque [2].

## 2.3 Vibro-acoustique et évanescence

En NAH, la prise de mesure se fait dans le champ proche de la source, qui se caractérise par la présence d'ondes évanescentes. Les ondes évanescentes sont des ondes à hautes fréquences spatiales dont l'amplitude décroit exponentiellement en s'éloignant orthogonalement de la source. Puisque la résolution du signal est bornée par la plus petite longueur d'onde du champ sonore enregistré, la contribution de ces ondes évanescentes au champ sonore est importante, car elle permet d'en améliorer la résolution.

On suppose ici que la source sonore est une plaque, que l'excitation de l'air est produite par des ondes de flexion et donc que la contribution des ondes longitudinales, de torsion et de cisaillement au champ sonore n'est pas significative. Physiquement, les ondes évanescentes sont une conséquence de la relation entre la fréquence angulaire $\omega$ et le nombre d'onde $k$, qui n'est pas la même pour l'onde de flexion et pour l'onde acoustique. En effet, les ondes de flexion sont dispersives, tandis que les ondes acoustiques ne le sont pas. Lorsque la vitesse de l'onde de flexion dans la source est plus petite que la vitesse du son dans l'air, l'onde de flexion ne peut exciter l'air

à la fois à une fréquence temporelle $\omega$ et à la fréquence spatiale $k$ correspondante à celle de l'onde dans la plaque. Cette idée est illustrée à la Figure 2.2 pour une plaque infinie.

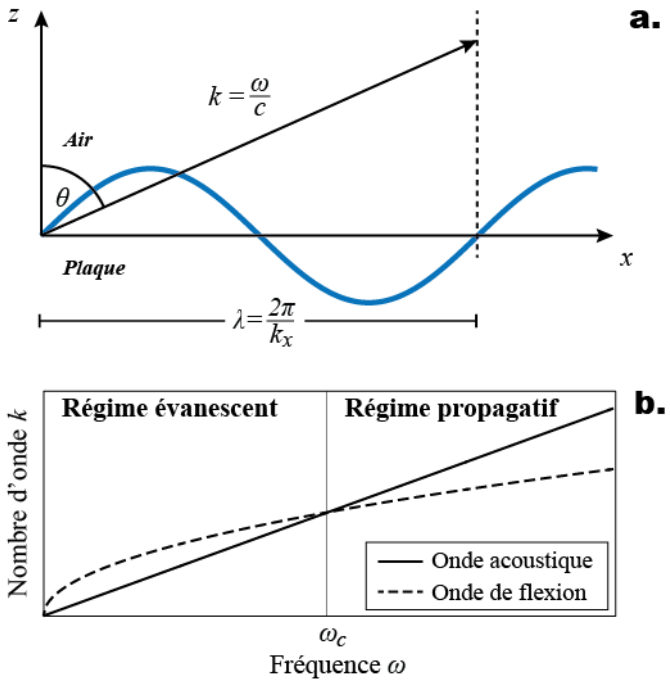


*Figure 2.2 : Rayonnement évanescent d'une plaque infinie. (a) Illustration de l'angle de rayonnement selon ω et k. (b) Nombre d'onde selon la fréquence pour une onde acoustique et une onde de flexion*

L'onde de flexion est illustrée en bleu à la Figure 2.2 (a). Celle-ci est caractérisée par une fréquence angulaire $\omega$ et une fréquence spatiale $k_x$ dans sa direction de propagation. L'excitation acoustique est produite par le déplacement transverse de la plaque qui génère un gradient de pression dans l'air : cela induit la propagation d'ondes acoustiques longitudinales dans l'air. Par conséquent, la fréquence temporelle de l'onde de flexion est la même que celle de l'onde acoustique. Similairement, la fréquence spatiale $k_x$ correspond à la projection du vecteur d'onde $k$ de l'onde acoustique sur le plan de la plaque. L'angle de rayonnement $\theta$ dépend donc de la relation entre $k$ et $\omega$ dans l'air et dans la plaque, où équivalemment, des relations de dispersion. Lorsqu'on l'exprime en fonction des vitesses de propagation, on retrouve la loi de Snell-Descartes [15, 16]:

$$\theta = \sin^{-1}\left(\frac{c}{c_f}\right), \tag{2.9}$$

où $c = \omega/k$ est la vitesse de l'onde acoustique et $c_f$ est la vitesse de l'onde de flexion dans la plaque :

$$c_f = \sqrt[4]{\frac{D}{\rho_p h}\omega^2}. \tag{2.10}$$

La variable $D$ représente la rigidité à la flexion de la plaque, $\rho_p$ sa densité et $h$ son épaisseur.

L'évanescence survient lorsque $c_f$ est plus petit que $c$. Dans ce cas, il n'y a pas de solution réelle à (2.9), et l'onde de flexion est réfléchie dans la plaque par réflexion totale interne. On pourrait alors percevoir le phénomène d'évanescence comme une manière d'outrepasser la discontinuité physique au point de réflexion. En effet, une telle discontinuité n'est pas une solution physique puisque le gradient de la pression à ce point tendrait vers l'infini, ce qui correspond à une énergie infinie ! Ainsi, l'onde rayonne dans l'air, mais s'atténue exponentiellement avant de devenir à nouveau une onde de flexion dans la plaque. L'atténuation exponentielle n'est donc pas due à un mécanisme de dissipation d'énergie, mais bien un transfert entre les deux médiums.

À la Figure 2.2 (b), le nombre d'onde en fonction de la fréquence angulaire est présenté pour l'onde de flexion et l'onde acoustique. La pente pour l'onde acoustique est constante, puisque l'onde est non-dispersive. La relation de dispersion pour l'onde de flexion suit une loi de racine carrée. Les deux courbes s'intersectent à la fréquence critique $\omega_c$ qui définit la limite entre les régimes évanescents ($\omega < \omega_c$) et propagatifs ($\omega > \omega_c$).

La condition d'évanescence est également mise en évidence par la fonction de Green exprimée dans l'espace de Fourier (Éq. (2.6)). L'apparition d'ondes évanescentes se concrétise lorsque la condition suivante est atteinte [2]:

$$k_z = \sqrt{k^2 - {k_x}^2 - {k_y}^2} \rightarrow k^2 < {k_x}^2 + {k_y}^2 \tag{2.11}$$

En effet, si la condition est atteinte, l'argument de la racine carrée est négatif et donc $k_z$ devient imaginaire. Par conséquent, la fonction de Green devient une exponentielle décroissante, par opposition à l'exponentielle complexe des ondes propagatives sinusoïdales. Le phénomène d'évanescence est donc intrinsèque à l'équation d'onde, et se concrétise lorsque le spectre de la

source comporte des composantes évanescentes, ou bien lorsqu'un bruit additif génère de telles composantes.

## 2.4 Formulation discrète et finie

En pratique, il n'est pas possible de faire l'acquisition du champ sonore sur toute la surface $D$ de rayon infini tel que développé à la section 2.1. Typiquement, on utilise une matrice rectangulaire de $N_x$ par $N_y$ microphones distribués uniformément. Les microphones sont distancés d'un pas constant $\Delta x = \Delta y$. Un schéma illustrant la matrice microphonique et les plans d'acquisition et de propagation est présenté à la Figure 2.3.

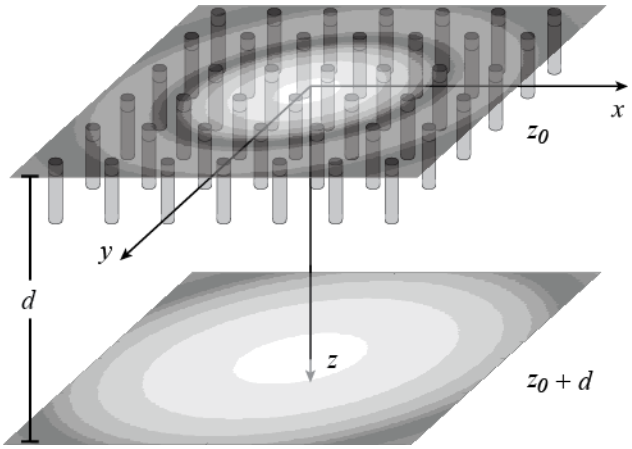


*Figure 2.3 : Plan d'acquisition (et matrice de microphones) et plan de propagation*

Le plan d'acquisition $z = z_0$ et le plan de propagation $z = z_0 + d$ sont centrés en $x = y = 0$, et la distance de propagation $d$ est la distance en $z$ entre les deux plans. Dans le domaine temporel, l'acquisition est également finie et discrète avec un pas d'acquisition $\Delta t$ et un nombre de points $N_t$.

L'échantillonnage sur un domaine fini est équivalent à multiplier la fonction analytique par un peigne de Dirac ainsi qu'une fenêtre rectangulaire 3-D :

$$\hat{p} = p(x, y, t) \cdot \text{Ш}_{xyt} \cdot \Pi_{xyt}, \tag{2.12}$$

avec le peigne de Dirac tridimensionnel $\text{Ш}_{xyt}$ :

$$Ш_{xyt} = \sum_{n_x=-\infty}^{\infty} \sum_{n_y=-\infty}^{\infty} \sum_{n_t=-\infty}^{\infty} \delta(x - \Delta x \cdot n_x) \cdot \delta\big(y - \Delta y \cdot n_y\big) \cdot \delta(t - \Delta t \cdot n_t), \qquad (2.13)$$

et la fenêtre tridimensionnelle $\Pi_{xyt}$ :

$$\Pi_{xyt} = \begin{cases} 1 & si\ n_\xi \in \left\{-\frac{N_\xi}{2}, \dots, \frac{N_\xi}{2}\right\} \cdot \Delta\xi, \\ 0 & ailleurs \end{cases} \qquad \forall \xi \in x, y, t. \qquad (2.14)$$

Tout dépendant de la formulation choisie, la fonction de Green est échantillonnée dans le domaine de Fourier ou dans le domaine direct. Ces multiplications avec Ш et Π s'appliquent donc à la fois au champ mesuré et à la fonction de transfert et introduisent des erreurs de traitement du signal dans le domaine direct ainsi que dans le domaine réciproque.

## 2.5 Convolution discrète

La convolution discrète se calcule normalement à partir de la transformée de Fourier discrète (*Discrete Fourier transform – DFT*). La *DFT* et son inverse sont définis aux équations (2.15) et (2.16) :

$$DFT(f[n]) = f[k] = \sum_{n=0}^{N-1} f[n] \cdot e^{-2\pi i n k/N}, \qquad (2.15)$$

$$DFT^{-1}(f[k]) = f[n] = \sum_{k=0}^{N-1} f[k] \cdot e^{2\pi i n k/N}, \qquad (2.16)$$

où $f[n]$ est un signal discret de $N$ éléments et $f[k]$ est le signal exprimé dans le domaine de Fourier discret, avec $n, k \in \{0, .., N-1\}$. En pratique, l'implémentation de la *DFT* se fait avec l'algorithme de transformée de Fourier rapide (*Fast Fourier transform – FFT*) qui permet de réduire le nombre d'opérations d'environ $N^2$ à $N \cdot log_2(N)$ [17].

Il a été établi dans les sections précédentes que le NAH consiste en la résolution d'une intégrale de convolution entre un champ de pression surfacique $p$ et une fonction de transfert $g$. Pour des fins de simplifications, la convolution 1-D est considérée dans cette section, même si en finalité, la convolution s'effectue selon trois dimensions en NAH.

La définition analytique de la convolution entre deux fonctions $p$ et $g$ est définie à l'équation (A.5), où $p$ et $g$ sont commutatifs. Lorsque les signaux sont définis sur un domaine fini, l'équation (A.5) ne peut être résolue puisqu'elle requiert une intégration sur $\mathbb{R}$. Il faut alors poser l'une des deux hypothèses suivantes, qui mènent chacune à un type de convolution discrète :

1. Convolution circulaire : les signaux finis $p$ et $g$ se répètent périodiquement de $-\infty$ à $+\infty$,
2. Convolution linéaire : les signaux $p$ et $g$ sont nuls pour toutes valeurs à l'extérieur du domaine de définition.

### 2.5.1 Convolution circulaire

La convolution circulaire s'exprime mathématiquement tel que présenté à l'équation (2.17) [5]. On suppose que les vecteurs $p$ et $g$ sont de même longueur $N$.

$$(p * g_N)(n)_{circ} \equiv \sum_{m=0}^{N-1} \left( \sum_{-\infty}^{\infty} p[m + kN] \right) g_N[n - m] \tag{2.17}$$

Il en résulte une suite périodique d'un signal de convolution de longueur $N$. Un exemple de convolution circulaire est présenté à la Figure 2.4.

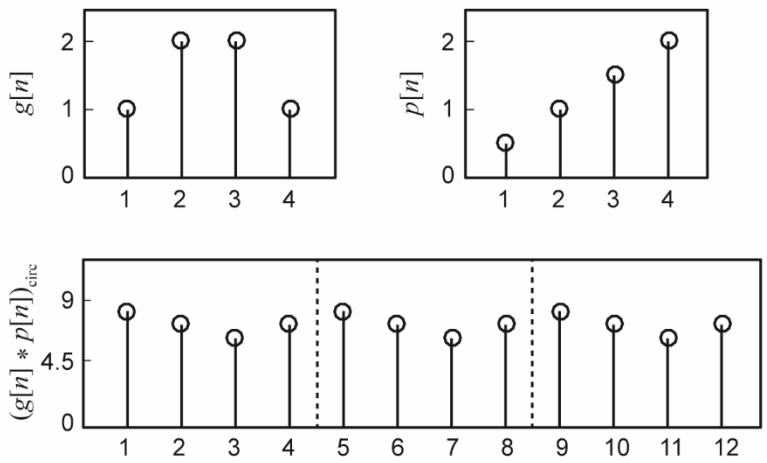


*Figure 2.4 : Exemple de convolution circulaire*

Malgré que la convolution circulaire soit constituée d'une succession infinie de répliques périodiques, elle est généralement représentée par une seule période de $N$ points. Une façon de

percevoir cette propriété de périodicité est d'exprimer la convolution circulaire sous forme matricielle [18, 19] :

$$h = (p * g_N)(n)_{circ} = P \cdot \vec{g} = \begin{bmatrix} p_1 & p_2 & p_3 & \cdots & p_N \\ p_N & p_1 & p_2 & & p_{N-1} \\ & \vdots & & \ddots & \vdots \\ p_2 & p_3 & p_4 & \cdots & p_1 \end{bmatrix}^T \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_N \end{bmatrix} \tag{2.18}$$

On remarque que la matrice *P* est circulante [19]. Les matrices circulantes sont des matrices de Toeplitz (diagonale constante) carrées pour lesquelles les éléments de chaque rangée sont identiques à ceux de la rangée précédente avec chaque élément translaté d'une position vers la droite. Le dernier élément de la rangée précédente se retrouve au début de la rangée subséquente, d'où la périodicité. Cela implique qu'il en va de même pour les colonnes. Une matrice circulante est donc définie par les $N$ éléments d'une rangée ou colonne. Pour la convolution circulaire, ceux-ci correspondent aux $N$ valeurs d'un des signaux à convoluer.

La matrice *P* est généralement non singulière et peut donc être diagonalisée avec son équation aux valeurs et vecteurs propres [18] :

$$FPF^{-1} = \Lambda = diag(\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_N), \tag{2.19}$$

où $F$ est une matrice contenant les $N$ vecteurs propres de la matrice circulante *P* et $\Lambda$ une matrice diagonale composée des $N$ valeurs propres. Il est intéressant de constater que pour toutes matrices circulante, $F$ est en réalité l'opérateur matriciel équivalent à la *DFT*, ce qui concorde avec la propriété de convolution de la transformée de Fourier. L'inverse de $F$ est équivalente à sa transposée. Les valeurs propres $\lambda$ correspondent aux $N$ valeurs de la transformée de Fourier discrète du signal $p_n$. Les valeurs explicites de $F$ et $\Lambda$ sont présentées aux équations (2.20) et (2.22).

$$F = \begin{bmatrix} (w^0) & (w^0) & (w^0) & & (w^0) \\ (w^0) & (w^1) & (w^2) & \cdots & \left(w^{(N-1)}\right) \\ (w^0) & (w^2) & (w^4) & & \left(w^{2(N-1)}\right) \\ & \vdots & & \ddots & \vdots \\ (w^0) & (w^{N-1}) & \left(w^{2(N-1)}\right) & \cdots & \left(w^{(N-1)(N-1)}\right) \end{bmatrix}, \tag{2.20}$$

Où $w$ est une fonction exponentielle complexe telle que définie à l'équation (2.21) :

$$w = e^{\frac{-2\pi i}{N}}, \tag{2.21}$$

et $\Lambda$, la matrice des valeurs propres :

$$\Lambda(k) = \begin{bmatrix} \lambda(1) & 0 & 0 & \dots & 0 \\ 0 & \lambda(2) & 0 & \dots & 0 \\ 0 & 0 & \lambda(3) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda(N) \end{bmatrix}, \tag{2.22}$$

avec,

$$\lambda(k) = \sum_{n=0}^{N-1} p_n e^{-2\pi i k n/N} = DFT(p_n). \tag{2.23}$$

En effet, on retrouve la propriété de convolution :

$$P = F^{-1}\Lambda F \tag{2.24}$$

En insérant (2.24) dans (2.18), on obtient;

$$h = F^{-1}\Lambda F g. \tag{2.25}$$

En multipliant de chaque côté par $F$ :

$$Fh = \Lambda F g, \tag{2.26}$$

et comme $F$ correspond à l'opérateur *DFT* et $\Lambda$ correspond à la *DFT* de $p$, on a que [18]:

$$DFT(h) = DFT(p) \cdot DFT(g). \tag{2.27}$$

### 2.5.2 **Convolution linéaire**

La convolution suppose que les vecteurs sont nuls à l'extérieur de leur domaine de définition. L'équation (2.28) présente la convolution linéaire sous forme de sommation :

$$(p * g_N)(n)_{lin} \equiv \sum_{m=0}^{N-1} p[n] \cdot g[n-m] \tag{2.28}$$

Où les valeurs de $p$ et $g$ sont nulles pour $m < 0$ et $m > N - 1$.

Lorsque la convolution linéaire est appliquée à deux vecteurs de longueur $N$, il en résulte un vecteur de $2N - 1$ points, concaténé à une séquence infinie de valeurs nulles. Un exemple de convolution linéaire est présenté à la Figure 2.5 :

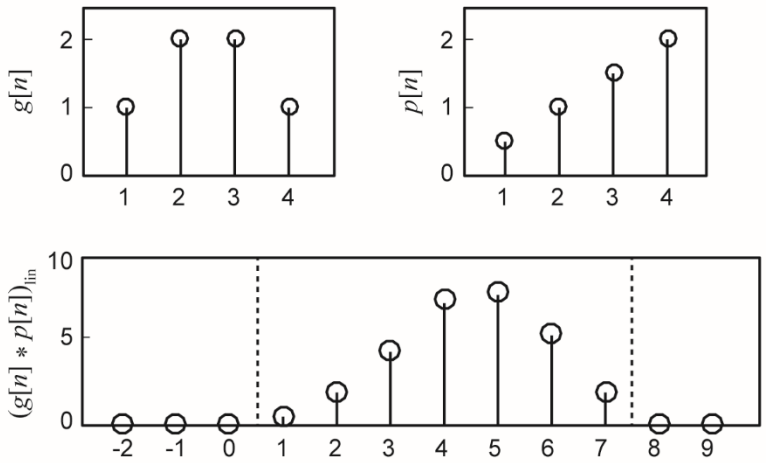


*Figure 2.5 : Exemple de convolution linéaire*

On remarque que les valeurs sont nulles à l'extérieur de $n \in \{1, \dots, N-1\}$, c'est-à-dire à l'extérieur de la zone délimitée par les lignes pointillées à la Figure 2.5.

L'équation (2.28) s'exprime également sous forme matricielle selon l'équation (2.29) :

$$(f * g_N)(n)_{lin} = F \cdot \vec{g} = \begin{bmatrix} f_1 & f_2 & f_3 & \cdots & f_N \\ 0 & f_1 & f_2 & & f_{N-1} \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & f_1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ f_N & 0 & 0 & & 0 \\ \vdots & & & \ddots & \vdots \\ f_2 & f_3 & f_4 & \cdots & 0 \end{bmatrix}^T \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_N \end{bmatrix}. \tag{2.29}$$

La matrice de convolution *F* est Toeplitz et est constituée de deux sous-matrices *N* par *N*. La somme de ces deux sous-matrices correspond à la matrice circulante de la convolution circulaire. Cependant, cette matrice Toeplitz n'est pas diagonalisée par la DFT comme c'est le cas pour la matrice circulante.

On peut cependant calculer la convolution linéaire avec la DFT en doublant les signaux convolués $p$ et $g$ par des valeurs nulles (un *zero-padding*) et en y appliquant la convolution circulaire : les répliques périodiques se multiplient alors avec les valeurs nulles, ce qui annule leur contribution :

$$(f * g_N)(n)_{lin} =$$

$$\left[\begin{array}{cccccc|cccccc}
f_1 & f_2 & f_3 & \cdots & f_{N-1} & f_N & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & f_1 & f_2 & \cdots & f_{N-2} & f_{N-1} & f_N & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & f_1 & \cdots & f_{N-3} & f_{N-2} & f_{N-1} & f_N & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 & \vdots \\
0 & 0 & 0 & \cdots & f_1 & f_2 & f_3 & f_4 & f_5 & \cdots & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & f_1 & f_2 & f_3 & f_4 & \cdots & f_N & 0 \\
\hline
0 & 0 & 0 & \cdots & 0 & 0 & f_1 & f_2 & f_3 & \cdots & f_{N-1} & f_N \\
f_N & 0 & 0 & \cdots & 0 & 0 & 0 & f_1 & f_2 & \cdots & f_{N-2} & f_{N-1} \\
f_{N-1} & f_N & 0 & \cdots & 0 & 0 & 0 & 0 & f_1 & \cdots & f_{N-3} & f_{N-2} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
f_3 & f_4 & f_5 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & f_1 & f_2 \\
f_2 & f_3 & f_4 & \cdots & f_N & 0 & 0 & 0 & 0 & \cdots & 0 & f_1
\end{array}\right]^T \left[\begin{array}{c} g_1 \\ g_2 \\ g_3 \\ \vdots \\ g_{N-1} \\ g_N \\ \hline 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{array}\right]. \qquad (2.30)$$

Toutefois, cela augmente d'un facteur deux le nombre de données à traiter et d'un facteur $2^3$ pour une convolution en 3-D [10]. On remarque que toutes les valeurs en dessous de la ligne horizontale dans le matrice de l'équation (2.30) s'annulent puisqu'elles sont multipliées par le *zero-padding* du vecteur $g$. Le résultat est donc exactement le même que celui de l'équation (2.29).

### 2.5.3 **Convolution modulée**

On effectue la convolution modulée en appliquant sur chaque signal une transformée de Fourier discrète généralisée (GDFT) et en multipliant les résultats avant d'appliquer une GDFT inverse. La GDFT et son inverse sont définies aux équations (2.31) et (2.32) [20, 21]:

$$MDFT(f_n) = \sum_{n=0}^{N-1} e^{-i\theta n/N} f_n e^{-2\pi i k n/N} = DFT\left(e^{-i\theta n/N} \cdot f_n\right), \qquad (2.31)$$

$$MDFT^{-1}(F_k) = \sum_{n=0}^{N-1} e^{i\theta n/N} F_k e^{2\pi i k n/N} = e^{i\theta n/N} DFT^{-1}(F_k). \qquad (2.32)$$

Le paramètre $\theta$ est un angle de modulation entre $-\pi$ et $\pi$. Pour $\theta = 0$, la GDFT est équivalente à la DFT. La DFT est donc un cas spécifique de la GDFT. La convolution modulée de deux vecteurs $h$ et $g$ s'exprime donc comme :

$$(h * g)(n)_{mod} = e^{i\theta n/N} DFT^{-1}\left(DFT\left(e^{-i\theta n/N} \cdot h\right) \cdot DFT\left(e^{-i\theta n/N} \cdot g\right)\right). \tag{2.33}$$

Cette équation est également associée à une formulation matricielle présentée à l'équation (2.34) :

$$(h * g_N)(n)_{mod} = H \cdot \vec{g} = \begin{bmatrix} h_1 & h_2 & h_3 & \cdots & h_N \\ \varphi h_N & h_1 & h_2 & & h_{N-1} \\ & \vdots & & \ddots & \vdots \\ \varphi h_2 & \varphi h_3 & \varphi h_4 & \cdots & h_1 \end{bmatrix}^T \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_N \end{bmatrix}, \tag{2.34}$$

$$\varphi = e^{i\theta/N}.$$

On remarque que l'expression matricielle est équivalente à celle de la convolution circulaire, avec les valeurs en dessous de la diagonale multipliées par une constante $\varphi$ qui dépend de l'angle de modulation $\theta$. De plus, la matrice $H$ s'exprime comme la somme des deux sous matrices de la convolution linéaire avec l'une de ces deux sous matrices étant modulée par la constante $e^{i\theta/N}$. Ainsi, on établit la relation suivante entre la convolution modulée et la convolution linéaire :

$$\begin{gathered}(p * g_N)(n)_{mod} = (p * g_N)(n)_{lin} + e^{i\theta}(p * g_N)(n+N)_{lin} \\ n \in \{0, \dots, N-1\}\end{gathered} \tag{2.35}$$

En définissant deux angles de modulation $\theta$ et en déterminant les deux convolutions modulées résultantes, l'équation (2.35) s'exprime comme un système de deux équations à deux inconnues, permettant de déterminer chaque moitié de la convolution linéaire [20, 21].

Si les signaux convolués sont réels, on peut résoudre le système avec une seule convolution modulée en choisissant $\theta$ tel que $e^{i\theta} \in \mathbb{C}$. Cette approche permet de réduire significativement les coûts de calcul pour la convolution linéaire, puisqu'elle ne nécessite pas l'utilisation d'une séquence de zéros (*zero-padding*).

De plus, en sélectionnant $\theta \to i\infty$, la modulation $e^{i\theta}$ tend vers zéro, ce qui permet d'obtenir les $N$ premières valeurs de la convolution linéaire avec un coût comparable à celui de la convolution

circulaire. En prenant une modulation réelle, on obtient une formulation équivalente à la transformée de Laplace numérique (*Numerical Laplace transform* – NLT).

## 2.6 Problèmes de traitement du signal

Trois principaux problèmes surviennent lors de la discrétisation des signaux sur des domaines finis. Ces problèmes causent l'écart entre les résultats théoriques de la formulation de Green et les résultats obtenus avec le NAH. Un principal défi en NAH est de sélectionner une formulation et des paramètres qui limitent ces erreurs.

### 2.6.1 Fuites spectrales

Les fuites spectrales (*leakage*) sont des erreurs qui s'expriment dans le domaine réciproque et qui sont causées par l'application de la DFT au signal fini et discret. Supposons le signal analytique $f(t)$ et $f(\omega)$ qui est son équivalent dans le domaine de Fourier. On obtient $f[n]$ en multipliant $f(t)$ par un peigne de Dirac et une fenêtre rectangulaire tel qu'exprimé à l'équation (2.12). On obtient $f[k]$ en prenant la DFT de $f[n]$. Les fuites spectrales s'expriment comme la différence entre $f[k]$ et la convolution de $f(\omega)$ avec la TF de la fenêtre et du peigne $(f(\omega) * \text{Ш}_\omega * \Pi_\omega)$. Si $\Delta t$ tend vers zéro et que $N_t$ tend vers l'infini, $f[k]$ et $(f(\omega) * \text{Ш}_\omega * \Pi_\omega)$ sont égaux.

On considère le peigne de Dirac 1-D suivant :

$$\text{Ш}_t = \sum_{n=-\infty}^{\infty} \delta(t - \Delta t \cdot n). \tag{2.36}$$

La transformée de Fourier de $\text{Ш}_t$ est :

$$\text{Ш}_\omega = \sum_{k=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi}{\Delta t} \cdot k\right). \tag{2.37}$$

où $\Delta\omega = \frac{2\pi}{\Delta t}$. Il y a une dualité entre le domaine direct et le domaine de Fourier : l'écart entre les pics de Dirac est inversement proportionnel d'un domaine à l'autre.

La multiplication de $f(t)$ par $\text{Ш}_t$ est équivalente à la convolution de $f(\omega)$ avec $\text{Ш}_\omega$ dans le domaine de Fourier. Cette convolution génère des répliques du spectre qui sont espacées de $\frac{2\pi}{\Delta t}$ :

$$f(\omega) * \text{Ш}_{\omega} = \sum_{k=-\infty}^{\infty} f\left(\omega - \frac{2\pi}{\Delta t} \cdot k\right). \tag{2.38}$$

En pratique, la fonction $f(\omega)$ est bornée entre $[-\omega_{max}, \omega_{max}]$. Ainsi, lorsque $\Delta t$ est suffisamment petit $\left(\frac{\pi}{\Delta t} > \omega_{max}\right)$, le théorème de Shannon-Nyquist est respecté et les répliques ne s'entrecroisent pas [22]. Autrement, il y a du repliement spectral (*aliasing*) et certaines composantes spectrales sont erronées.

De manière similaire, la multiplication avec la fenêtre rectangulaire $\Pi_t$ génère des fuites dans le domaine réciproque. La transformée de Fourier de la fonction rectangulaire est un sinus cardinal :

$$\mathcal{F}\{\Pi_t\} = \frac{\sin(\pi t/\tau)}{\pi t/\tau}, \tag{2.39}$$

où $\tau$ est la largeur de la fenêtre. La convolution avec la fonction sinus cardinal génère des composantes erronées qui s'atténuent à mesure que $\tau$ devient grand.

En NAH, puisque les sources étudiées rayonnent sur un domaine spectral fini, on choisit une fréquence d'acquisition suffisamment grande et un espace inter-microphonique suffisamment petit pour éviter la génération de repliement spectral sur le champ mesuré. Cependant, lorsque la fonction de Green est échantillonnée dans le domaine de Fourier, des fuites spectrales sont générées puisque le champ propagé est représenté dans le domaine temps-espace. Certains auteurs proposent l'utilisation de fenêtres pour amoindrir l'effet des fuites spectrales [2, 6], cependant, ces méthodes sont imparfaites, car elles affectent également l'amplitude de la solution.

### 2.6.2 **Recouvrement**

Les erreurs de recouvrement (*wrap-around*) sont dues à l'application d'une convolution circulaire lorsque la convolution linéaire aurait dû être appliquée. En d'autres mots, il s'agit de l'erreur occasionnée par la supposition que le problème physique comporte une périodicité spatiale ou temporelle, alors qu'en réalité, les valeurs sont nulles à l'extérieur du domaine de convolution. En se rapportant aux exemples de convolutions circulaires et linéaires présentés aux Figure 2.4 et Figure 2.5, il est clair que les résultats sont significativement différents.

L'erreur de *wrap-around* est courante en NAH puisqu'on traite le problème dans l'espace de Fourier, et ce malgré que l'amplitude du champ sonore tende vers zéro en s'éloignant de la source dans le domaine spatial. Cette erreur devient d'autant plus évidente lorsqu'on traite des problèmes non-stationnaires causaux puisque le *wrap-around* contrevient aux relations de causalité du champ. En effet, le champ doit rester nul jusqu'à ce que l'onde ait atteint la position d'observation.

Certains auteurs proposent d'appliquer un *zero-padding* au champ mesuré [2, 8], toutefois, cela ne fait qu'amoindrir le phénomène. Pour obtenir la convolution linéaire et supprimer complètement l'erreur de recouvrement, le *zero-padding* doit être appliqué à la mesure ainsi qu'à la fonction de Green. Cependant, cela n'est pas possible si la fonction analytique est définie dans l'espace de Fourier. Finalement, lorsque la source sonore est périodique dans le temps, l'hypothèse de périodicité de la convolution est adéquate et désirable : il faut appliquer la convolution circulaire dans le domaine temporel.

### 2.6.3 Troncature

Malgré que l'amplitude de la source sonore converge vers zéro, son acquisition sur un domaine fini engendre parfois une troncature du champ. Cette troncature peut survenir dans le domaine temporel ou dans le domaine spatial. Les composantes tronquées contribuent également au champ calculé, cependant, leur amplitude est supposée nulle lorsqu'on applique la convolution linéaire. L'erreur de troncature correspond donc à ces contributions manquantes. Ainsi, plus la fenêtre de mesure est large, moins les erreurs de troncature sont importantes. Les erreurs de troncature sont également issues du fenêtrage de la fonction de Green.

## 2.7 Déconvolution discrète

Afin de déterminer le champ sonore entre le plan d'acquisition et la source, le problème doit être inversé. Le plan de propagation devient alors le plan de mesure et le plan de mesure devient le plan sur lequel on détermine le champ sonore. La manière de résoudre ce problème est de substituer la convolution par une déconvolution, son opérateur inverse. On appelle alors cette opération inverse la « rétro-propagation ». La géométrie du problème est illustrée à la Figure 2.6.

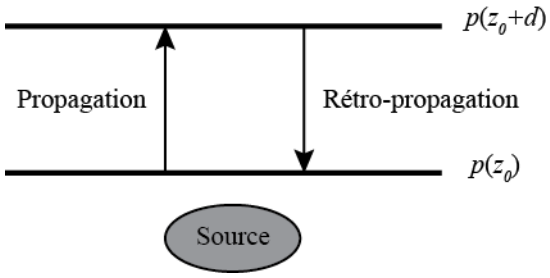


*Figure 2.6 : Géométries du problème direct et inverse*

On peut calculer la déconvolution en exprimant le problème fini et discret sous forme matricielle :

$$p(z_0 + d) = G \cdot p(z_0), \tag{2.40}$$

où $p$ est le champ de pression aux plans indiqués et $G$ est la fonction de Green exprimée sous forme de matrice de convolution. On peut ainsi déterminer $p(z_0)$ à partir de $p(z_0 + d)$ en inversant la matrice de convolution $G$ :

$$p(z_0) = G^{-1} \cdot p(z_0 + d). \tag{2.41}$$

$G^{-1}$ est la matrice de déconvolution.

Puisqu'en NAH, la convolution s'exprime en trois dimensions, la matrice de convolution $G$ contient un nombre d'éléments équivalent à $\left(N_t N_x N_y\right)^2$. Afin de respecter la condition de Nyquist et minimiser les erreurs de troncature, $N$ est au moins de l'ordre de $10^2$ pour chaque dimension. Ainsi, la matrice $G$ contient au moins $10^{12}$ éléments, ce qui rend le calcul de son inverse trop coûteux.

En supposant la convolution circulaire (ou la convolution linéaire avec du remplissage à zéro), la matrice de convolution $G$ prend la forme d'une matrice block-circulante telle que présentée à l'équation (2.42).

$$G = \begin{bmatrix} \begin{bmatrix} c_1^1 & c_2^1 & \cdots & c_{N_y}^1 \\ c_{N_y}^1 & c_1^1 & \cdots & c_2^1 \\ \vdots & \vdots & \ddots & \vdots \\ c_2^1 & c_3^1 & \cdots & c_1^1 \end{bmatrix}^T & \begin{bmatrix} c_1^2 & c_2^2 & \cdots & c_{N_y}^2 \\ c_{N_y}^2 & c_1^2 & \cdots & c_2^2 \\ \vdots & \vdots & \ddots & \vdots \\ c_2^2 & c_3^2 & \cdots & c_1^2 \end{bmatrix}^T & \cdots & \begin{bmatrix} c_1^{N_t} & c_2^{N_t} & \cdots & c_{N_y}^{N_t} \\ c_{N_y}^{N_t} & c_1^{N_t} & \cdots & c_2^{N_t} \\ \vdots & \vdots & \ddots & \vdots \\ c_2^{N_t} & c_3^{N_t} & \cdots & c_1^{N_t} \end{bmatrix}^T \\ \begin{bmatrix} c_1^{N_t} & c_2^{N_t} & \cdots & c_{N_y}^{N_t} \\ c_{N_y}^{N_t} & c_1^{N_t} & \cdots & c_2^{N_t} \\ \vdots & \vdots & \ddots & \vdots \\ c_2^{N_t} & c_3^{N_t} & \cdots & c_1^{N_t} \end{bmatrix}^T & \begin{bmatrix} c_1^1 & c_2^1 & \cdots & c_{N_y}^1 \\ c_{N_y}^1 & c_1^1 & \cdots & c_2^1 \\ \vdots & \vdots & \ddots & \vdots \\ c_2^1 & c_3^1 & \cdots & c_1^1 \end{bmatrix}^T & \cdots & \begin{bmatrix} c_1^{N_t-1} & c_2^{N_t-1} & \cdots & c_{N_y}^{N_t-1} \\ c_{N_y}^{N_t-1} & c_1^{N_t-1} & \cdots & c_2^{N_t-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_2^{N_t-1} & c_3^{N_t-1} & \cdots & c_1^{N_t-1} \end{bmatrix}^T \\ \vdots & \vdots & \ddots & \vdots \\ \begin{bmatrix} c_1^2 & c_2^2 & \cdots & c_{N_y}^2 \\ c_{N_y}^2 & c_1^2 & \cdots & c_2^2 \\ \vdots & \vdots & \ddots & \vdots \\ c_2^2 & c_3^2 & \cdots & c_1^2 \end{bmatrix}^T & \begin{bmatrix} c_1^3 & c_2^3 & \cdots & c_{N_y}^3 \\ c_{N_y}^3 & c_1^3 & \cdots & c_2^3 \\ \vdots & \vdots & \ddots & \vdots \\ c_2^3 & c_3^3 & \cdots & c_1^3 \end{bmatrix}^T & \cdots & \begin{bmatrix} c_1^1 & c_2^1 & \cdots & c_{N_y}^1 \\ c_{N_y}^1 & c_1^1 & \cdots & c_2^1 \\ \vdots & \vdots & \ddots & \vdots \\ c_2^1 & c_3^1 & \cdots & c_1^1 \end{bmatrix}^T \end{bmatrix}^T, \quad (2.42)$$

avec

$$c_i^j = \begin{bmatrix} g\big(n_x[1], n_y[i], n_t[j]\big) & g\big(n_x[2], n_y[i], n_t[j]\big) & \cdots & g\big(n_x[N_x], n_y[i], n_t[j]\big) \\ g\big(n_x[N_x], n_y[i], n_t[j]\big) & g\big(n_x[1], n_y[i], n_t[j]\big) & \cdots & g\big(n_x[N_x - 1], n_y[i], n_t[j]\big) \\ \vdots & \vdots & \ddots & \vdots \\ g\big(n_x[2], n_y[i], n_t[j]\big) & g\big(n_x[3], n_y[i], n_t[j]\big) & \cdots & g\big(n_x[1], n_y[i], n_t[j]\big) \end{bmatrix}^T. \quad (2.43)$$

Le vecteur $p(z_0 + d)$ de l'équation (2.41) correspond à la concaténation selon la première dimension des vecteurs de la matrice d'acquisition 3-D. L'inverse de la matrice block-circulante est diagonalisée par l'opérateur DFT 3-D conformément à l'équation (2.24) appliquée au cas 1-D. Le champ rétro-propagé est donc obtenu en divisant la DFT 3-D de la mesure par la DFT 3-D de la fonction de Green :

$$DFT_{3D}\big(p(z_0)\big) = \frac{DFT_{3D}\big(p(z_0 + d)\big)}{DFT_{3D}(g)}. \quad (2.44)$$

## 2.8 Régularisation

La déconvolution est une opération mal posée dans le sens où elle est instable mathématiquement : un faible bruit ajouté à la mesure cause une erreur considérable. Ce concept est exposé à la Figure 2.7.

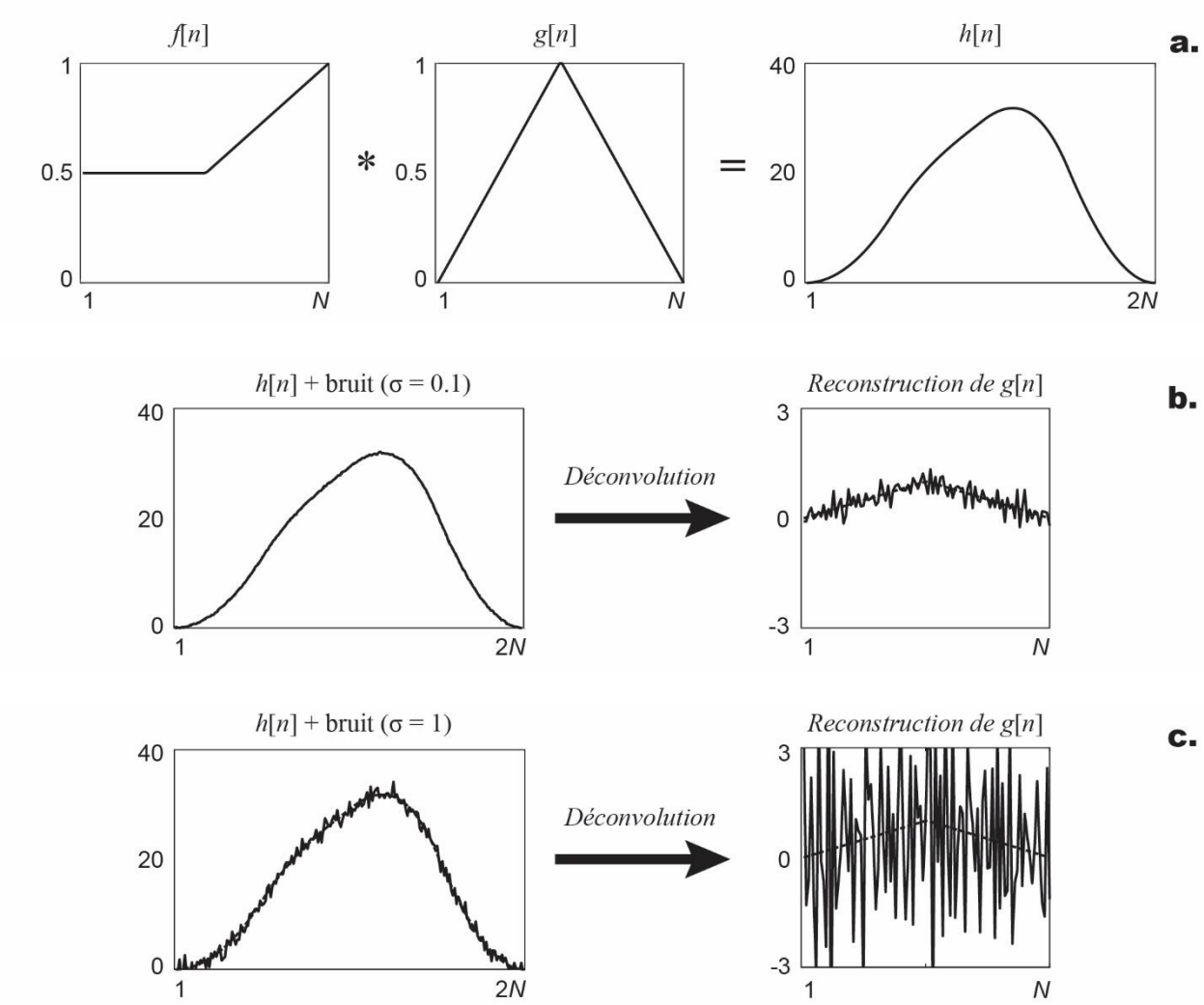


*Figure 2.7 : Exemple d'amplification du bruit causé par la déconvolution*

À la Figure 2.7 a), on considère deux signaux quelconques $f[n]$ et $g[n]$ qu'on convolue linéairement pour obtenir $h[t]$. À la Figure 2.7 b), on applique un bruit gaussien avec un écart-type $\sigma$ de 0.1 au résultat de convolution $h[n]$, qu'on considère ici comme la mesure. Ensuite, on applique la déconvolution à partir du signal $f[n]$ afin de retrouver $g[n]$ : malgré le faible bruit ajouté, la reconstruction est considérablement bruitée. À la Figure 2.7 c), on applique le même processus, mais en ajoutant un bruit ayant un écart-type dix fois plus grand ($\sigma = 1$), mais somme toute faible par rapport à l'amplitude de la mesure. La reconstruction obtenue est très mauvaise : l'amplitude de l'erreur est beaucoup plus grande que celle de la solution.

En pratique, les signaux convolués sont bornés fréquentiellement et en supposant la présence d'un bruit additif, les composantes hautes-fréquences de faible amplitude ont un faible ratio signal sur bruit. Puisque la déconvolution réamplifie plus fortement les composantes hautes fréquences, ces composantes bruitées deviennent dominantes dans le signal reconstruit.

En NAH, puisque le champ proche contient des ondes évanescentes qui décroissent exponentiellement en s'éloignant de la source, celles-ci sont réamplifiées lors de la rétro-propagation. Ce sont ces composantes évanescentes hautes fréquences dont l'amplitude est très faible au plan de mesure qui contribuent significativement à l'erreur lors de la rétro-propagation.

### 2.8.1 **Régularisation de Tikhonov**

Ainsi, il est nécessaire d'appliquer une forme de régularisation aux signaux afin de limiter la génération d'erreurs de rétro-propagation. La régularisation de Tikhonov est la méthode la plus répandue, puisqu'elle est facilement applicable et a de bonnes performances dans la majorité des cas. Elle repose sur l'utilisation d'une fonction coût entre la solution $p_s$ exprimée au plan de reconstruction, l'opérateur de convolution $G$ et la mesure $p_m$ :

$$\|Gp_s - p_m\|^2 + \lambda\|p_s\|^2, \tag{2.45}$$

où $\lambda$ est le coefficient de régularisation et $\|.\|$ correspond à la norme $\ell^2$. L'élément de droite de l'équation (2.45) correspond à la norme de la solution : lorsque le problème est peu régularisé, le bruit est fortement réamplifié et la norme de la solution domine l'équation. À l'opposé, lorsque le problème est trop régularisé, la solution $p_s$ tend vers zéro et le terme de résidu à gauche de l'équation (2.45) domine. Il y a donc un équilibre entre ces deux termes qu'on obtient en choisissant une pondération $\lambda$ adéquate et en minimisant (2.45) selon $p_s$.

On peut minimiser (2.45) analytiquement en mettant la dérivée par rapport à $p_s$ égale à zéro :

$$\frac{\partial}{\partial p_s}(G^2 p_s^2 - 2Gp_m p_s + p_m^2 + \lambda p_s^2) = 0. \tag{2.46}$$

La solution est :

$$\widehat{p_s} = \frac{G}{G^2 + \lambda} p_m = \frac{G^2}{G^2 + \lambda} p_s, \tag{2.47}$$

où $\widehat{p_s}$ est la solution régularisée. Lorsque $G$ est circulante (ou block-circulante), on peut interpréter l'équation (2.47) dans le domaine de Fourier, et le terme $G^2/(G^2 + \lambda)$ devient analogue à un filtre de Wiener : un filtre passe-bas dont la fréquence de coupure est $\lambda$ [9]. Le défi est de trouver la valeur du paramètre de régularisation $\lambda$ qui permet de filtrer adéquatement la reconstruction.

### 2.8.2 **Régularisation parcimonieuse**

Lorsque la solution $p_s$ contient une grande proportion de valeurs nulles, la régularisation de Tikhonov n'arrive pas à bien prédire ces valeurs. Ceci est dû au fait que le terme $\|p_s\|^2$ tend vers zéro pour des valeurs de $p_s$ plus petites que l'unité, malgré que ces valeurs ne soient pas nulles. Cela devient problématique dans la reconstruction de signaux causaux, puisqu'une faible amplitude va à l'encontre du principe de causalité qui implique une solution nulle pour plusieurs composantes.

Il devient alors préférable d'utiliser la fonction coût suivante :

$$argmin\{\|Gp_s - p_m\|^2 + \lambda|p_s|\}, \tag{2.48}$$

où $|.|$ représente la norme $\ell^1$. En effet, étant donné que le terme $|p_s|$ n'est pas au carré, celui-ci doit être très faible pour obtenir une bonne estimation de la véritable de solution égale à zéro [23].

Cependant, dans le cas de la régularisation parcimonieuse, il n'y a pas de solution analytique au problème. Il faut alors minimiser l'équation (2.48). Puisque l'équation est convexe, on peut y appliquer un algorithme d'optimisation convexe [24]. Cela implique l'utilisation d'algorithmes itératifs et le coût de calcul est augmenté significativement par rapport à la régularisation de Tikhonov.

## 2.9 Détermination du coefficient de régularisation

Plusieurs méthodes permettent d'évaluer la valeur optimale du paramètre de régularisation $\lambda_{opt}$. La performance de ces méthodes dépend à la fois de l'opérateur utilisé et des données mesurées : il est possible qu'une méthode soit performante pour une fonction de Green donnée, et le soit moins pour une autre ou pour des mesures ayant d'autres caractéristiques (parcimonie, bruit, etc.).

Trois méthodes sont présentées dans cette section. La valeur de $\lambda_{opt}$ dépend du rapport signal sur bruit de la mesure, cependant, les méthodes présentées ne nécessitent pas d'avoir une connaissance initiale de ce niveau de bruit.

### 2.9.1 **Courbe en L**

La méthode de la courbe en L consiste à représenter le logarithme de la norme du résidu $\|G \cdot \widehat{p_s} - p_m\|$ selon le logarithme de la norme de la solution $\|\widehat{p_s}\|$. Cette représentation est effectuée pour plusieurs valeurs de $\lambda$ : l'ensemble des valeurs génère une courbe ayant la forme d'un « L ». Un exemple de courbe en L est présenté à la Figure 2.8.

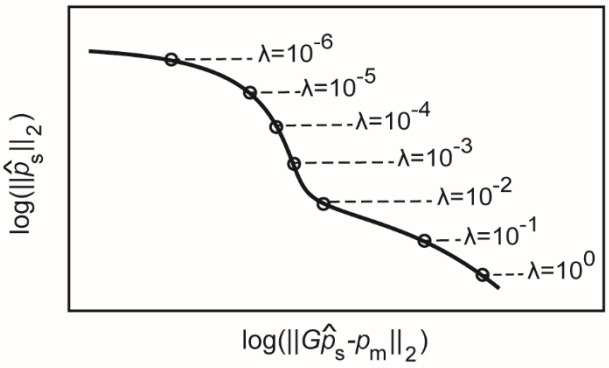


*Figure 2.8 : Exemple de courbe en L pour la régularisation de Tikhonov*

On suppose alors qu'un compromis existe entre les deux termes de la fonction coût et que celui-ci correspond à la valeur optimale du paramètre de régularisation $\lambda_{opt}$. Ce compromis correspond au point de courbure maximale de la courbe en L, qu'on peut déterminer numériquement [9, 25]. Dans l'exemple de la Figure 2.8, la prédiction de la valeur optimale du paramètre de régularisation serait entre $10^{-3}$ et $10^{-2}$.

La méthode de la courbe en L est également utilisée pour la régularisation parcimonieuse [26, 27]. On considère alors le logarithme de la norme $\ell^1$ de la solution en ordonnée.

### 2.9.2 **Validation croisée généralisée**

La validation croisée consiste à considérer le vecteur de mesure $p_m$ auquel on retire sa $n^{ième}$ donnée : on dénote alors ce vecteur $p_m^{[n]}$. Ensuite, on trouve la solution $p_s^{[n]}$ correspondante avec la méthode de Tikhonov pour une plage de valeurs de $\lambda$. On utilise ensuite $p_s^{[n]}$ et l'opérateur de

propagation $G(\lambda)$ pour évaluer $p_m$ au point où la donnée a été initialement retirée, ce qui permet de déterminer l'écart produit par le modèle pour un certain paramètre $\lambda$. On reproduit ensuite cette procédure en retirant tour à tour chaque valeur de $p_m$ et on calcule l'écart au sens des moindres carrés pour l'ensemble des cas. La fonction coût correspondante est :

$$V_n(\lambda) = \frac{1}{N}\sum_{n=1}^{N}\left(p_m^{[n]} - G(\lambda)\cdot p_s^{[n]}\right)^2. \tag{2.49}$$

Il suffit alors de minimiser (2.49) selon $\lambda$ pour trouver le paramètre de régularisation optimal, c'est-à-dire le propagateur $G(\lambda)$ qui produit la meilleure prédiction pour l'ensemble des points de mesure. Cette méthode de validation croisée est cependant très coûteuse en temps de calcul, puisqu'elle requiert de résoudre le problème pour chaque valeur de $\lambda$ et pour l'ensemble des données retirées.

La méthode de validation croisée généralisée produit un résultat presque équivalent, mais de manière beaucoup plus efficace, en se basant sur le lemme de la donnée manquante [28]. On obtient alors une fonction coût qui dépend de $\lambda$ [29]:

$$J_{GCV}(\lambda) = \frac{\|G\cdot\widehat{p_s} - p_m\|^2}{\left(Tr\left(I - \frac{|G|^2}{|G|^2+\lambda}\right)\right)^2}, \tag{2.50}$$

où $I$ est l'identité et $Tr(.)$ est l'opérateur trace. Il suffit alors de minimiser (2.50) pour évaluer $\lambda_{opt}$. Lorsque $G$ est circulante, on peut simplifier (2.50) en l'exprimant dans le domaine de Fourier en usant du théorème de Parseval [29].

### 2.9.3 **Approche Bayésienne**

L'approche Bayésienne consiste à considérer le problème inverse avec une approche probabiliste : la pression mesurée $p_m$ et la pression à la source $p_s$ sont interprétées comme des fonctions densité de probabilité (PDF) ayant une certaine distribution. Par exemple, le champ mesuré $p_m$ varie selon les diverses sources de bruit et une mesure répétée du champ de pression contiendra certaines variations statistiques. En prenant de multiples acquisitions, on peut déterminer la nature et les

paramètres de cette distribution, ou bien, on peut supposer qu'elle ait une certaine forme en se basant sur des hypothèses statistiques.

Les champs acoustiques au plan de mesure et au plan source sont mis en relation avec le théorème de Bayes :

$$P(p_s|p_m) \cdot P(p_m) = P(p_m|p_s) \cdot P(p_s). \tag{2.51}$$

Pour la résolution du problème inverse, on cherche à déterminer le maximum de la PDF a posteriori $P(p_s|p_m)$, qui correspond à la solution la plus probable pour le champ $p_s$ sachant avoir obtenu une mesure $p_m$. L'équation (2.51) devient donc :

$$P(p_s|p_m) = \frac{P(p_m|p_s) \cdot P(p_s)}{P(p_m)}, \tag{2.52}$$

où $P(p_m|p_s)$ est la fonction de vraisemblance qui contient l'information probabiliste sur la mesure et $P(p_s)$ est l'apriori, qui contient l'information initiale sur la source étudiée. Le terme $P(p_m)$ permet de normaliser l'équation de manière à ce que son intégrale soit unitaire, une propriété de toute PDF.

Conformément au théorème central limite, on choisit une forme gaussienne pour la fonction de vraisemblance, puisque le bruit de la mesure est généralement issu d'évènements indépendants. Avec l'approche préconisée par Pereira et Antoni, on ne suppose pas connaitre la variance, mais seulement la nature de la distribution [30, 31]. Pour l'apriori, malgré l'absence de connaissance sur la source, il est conservateur de choisir une distribution gaussienne, puisque c'est la distribution ayant la plus grande entropie de Shannon. La maximisation de l'équation (2.52) obtenue en prenant le logarithme des PDF gaussiennes mène alors à la fonction coût de Tikhonov (éq. (2.45)). Une force de l'approche Bayésienne est que le coefficient de régularisation $\lambda$ s'exprime comme $\beta^2/\alpha^2$, où $\beta^2$ est la variance de la PDF gaussienne de la fonction de vraisemblance et $\alpha^2$ est la variance de la PDF gaussienne de la source. On peut donc interpréter le coefficient de régularisation comme le ratio bruit sur signal.

Pereira et al. ont démontré qu'il est possible d'évaluer la valeur optimale de $\lambda$ en appliquant le théorème de Bayes sur le terme de normalisation $P(p_m|\alpha^2, \beta^2) = \int_{-\infty}^{\infty} P(p_m|p_s) \cdot P(p_s)\, dx$ [30]. On suppose des distributions pour $\alpha^2$ et $\beta^2$ et on obtient :

$$P(\alpha^2, \beta^2 | p_m) \propto P(p_m | \alpha^2, \beta^2) \cdot P(\alpha^2, \beta^2). \quad (2.53)$$

En supposant des distributions uniformes pour $P(\alpha^2, \beta^2)$ et en maximisant l'apostériori, on obtient une fonction coût qu'on minimise pour obtenir $\lambda_{opt}$ :

$$\mathrm{J}_{Bayes}(\lambda) = \sum ln(|\mathcal{F}\{g\}|^2 + \lambda) + N ln\left(\frac{1}{N}\sum \frac{|\mathcal{F}\{p_m\}|^2}{|\mathcal{F}\{g\}|^2 + \lambda}\right). \quad (2.54)$$

On suppose ici une forme circulante pour $g$, ce qui permet d'exprimer l'équation (2.54) dans le domaine de Fourier de manière équivalente à la formulation en valeurs singulières présentée dans [30].

Lorsque nous savons que la solution est parcimonieuse, un apriori Laplacien est utilisé puisqu'il favorise cette configuration [23]. Dans ce cas, le maximum apriori est équivalent à la minimisation de l'équation (2.48).

## 2.10 Méthodes d'holographie acoustique non-stationnaire

On présente dans ce sous-chapitre une revue de la littérature des différentes adaptations de l'holographie acoustique en champ proche appliquée à la détermination de champs rayonnés par des sources non-stationnaires. Une synthèse des forces et des faiblesses des différentes méthodes est également présentée. Ceci permet de contextualiser les objectifs de la thèse qui sont définis au chapitre suivant.

### 2.10.1 « Time Domain holography » et méthodes analogues

La méthode standard d'holographie acoustique consiste à faire l'acquisition du champ de pression auquel on applique une *TF* 3-D, avant de le multiplier avec la fonction de Green dans l'espace $\omega$-$k$. Puis, une *TF* spatiale inverse est appliquée pour représenter le champ de pression $p(\omega, z_0 + d)$ selon la fréquence. La méthode standard suppose que la source acoustique est stationnaire et périodique.

Hald a introduit le « *Time Domain Holography – TDH* » qui consiste à représenter le signal au plan propagé dans le domaine temporel $t$ au lieu du domaine fréquentiel $\omega$ en appliquant une *TF* temporelle inverse [32]. De cette manière, on peut visualiser l'évolution temporelle du signal au

plan d'observation, ce qui est pertinent dans le cas de signaux non-stationnaires puisque leur contenu fréquentiel change avec le temps.

À la Figure 2.9, une représentation schématisée des opérations du *TDH* est présentée.

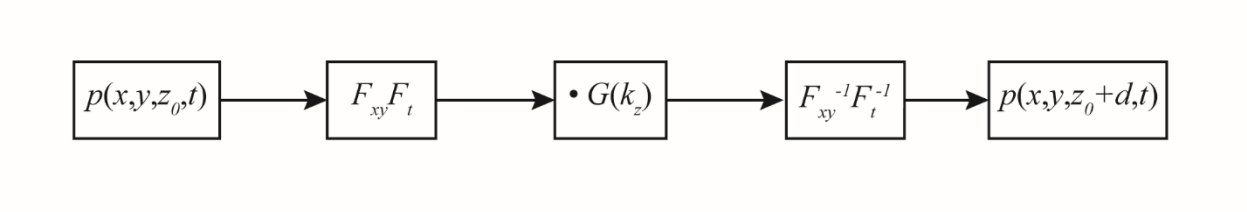


*Figure 2.9 : Schéma bloc des opérations mathématiques effectuées pour la méthode de « Time Domain Holography »*

Étant donnés la multiplication et l'échantillonnage de la fonction de Green dans le domaine des fréquences, cette méthode génère d'importantes erreurs de fuites spectrales. De plus, la convolution est circulante et du recouvrement est généré. L'impact du recouvrement peut être amoindri en ajoutant un remplissage à zéro au champ de pression mesuré, cependant, cela ne supprime pas complètement les répliques. Si la source étudiée ne rayonne pas avant un temps $t_0$, d'importantes erreurs de causalité résultent du phénomène de repliement. Deblauwe et al. ont également proposé une formulation similaire [33].

La Rochefoucauld et al. [7] ont analysé le *TDH* pour une propagation de signaux expérimentaux et simulés. L'analyse a été effectuée avec la formulation de Hald $(k_x, k_y, \omega)$, mais également en résolvant l'intégrale de convolution dans une combinaison de domaines fréquentiels et directs (les couples $(x, y, t)$ et $(x, y, \omega)$). Les simulations sont effectuées pour un piston au profil de vitesse gaussien. Le rayon du piston, la distance entre les microphones et la distance de propagation sont les principaux facteurs étudiés. De manière générale, la méthode effectuant la convolution effectuée dans le domaine de Fourier présente la plus petite erreur et requiert un temps de calcul plus court. Cependant, pour la grande distance de propagation, les méthodes ne résolvant pas la convolution dans l'espace de Fourier sont plus performantes. Dans le cadre de cette analyse, il n'est pas spécifié de quelle manière la convolution est calculée : il est probable que la convolution soit circulaire et que du repliement soit la principale cause d'erreur.

### 2.10.2 **Real-Time Near-field Acoustical Holography**

Pascal, Grulier et Thomas [6, 34] introduisent le « *Real-Time Near-field Acoustical Holography – RT-NAH* » dont l'objectif est de traiter la propagation de signaux non-stationnaires en continu. Contrairement au *TDH* qui requiert l'acquisition temporelle complète du signal pour calculer sa transformé de Fourier, le RT-NAH effectue la convolution dans le domaine temps – nombres d'onde $(k_x, k_y, t)$. Cette formulation s'interprète comme la convolution 1-D selon $t$ de chaque élément du spectre des nombres d'ondes avec la fonction de transfert.

Pour ce faire, une formulation alternative de la fonction de Green est utilisée, et est présentée à l'équation (*2.55*):

$$h(k_x, k_y, z, t) = \delta\left(t - \frac{z}{c}\right) - z\sqrt{{k_x}^2 + {k_y}^2}\,\frac{J_1\left(c\sqrt{{k_x}^2 + {k_y}^2}\sqrt{t^2 - \frac{z^2}{c^2}}\right)}{\sqrt{t^2 - \frac{z^2}{c^2}}}\,\Gamma\left(t - \frac{z}{c}\right), \qquad (2.55)$$

où $J_1$ est la fonction de Bessel de première espèce d'ordre un et $\Gamma$ est la fonction de Heaviside. Cette fonction est obtenue en résolvant l'intégrale de Helmholtz dans le domaine de Laplace [35].

Avec cette méthode, la convolution temporelle est linéaire, ce qui permet de supprimer les erreurs de recouvrement dans ce domaine. Cependant, comme l'opérateur est exprimé dans le domaine des nombres d'ondes, la convolution spatiale est circulaire, ce qui génère des erreurs de repliement. De plus, l'échantillonnage de la fonction de Green dans l'espace des $k$ génère des fuites spectrales dans le domaine direct.

Étant donné que la convolution est une opération linéaire, un avantage mis en évidence par Thomas et al. est que cette méthode permet de traiter des signaux non-stationnaires dont l'amplitude selon le temps ne converge pas rapidement vers zéro. En effet, en effectuant une combinaison linéaire du résultat de la convolution entre le propagateur et plusieurs acquisitions successives, il est possible de reconstruire le champ propagé par morceaux.

La régularisation des signaux rétro-propagés avec cette méthode est un aspect qui a été abordé dans les travaux de Paillasseur et al. [36], puis dans ceux de Le Magueresse et al. considérant une approche Bayésienne [37].

### 2.10.3 Time-Domain Plane-Wave Superposition Method (TD-PWSM)

Le *TD-PWSM* a été développée par Zhang, Pascal et Thomas [38]. Son objectif est de reconstituer le champ sonore à la source avec l'opérateur temps-nombre d'onde exprimé à l'équation (*2.55*). La déconvolution est résolue dans le domaine temporel pour chaque position et une combinaison linéaire des résultats est effectuée pour obtenir le champ en entier.

Cette déconvolution n'est pas calculée avec la transformée de Fourier rapide mais avec un terme de sommation, ce qui augmente significativement le temps de calcul selon les auteurs. Par contre, cela permet de changer la résolution spatiale en prenant des termes $\Delta k_x$ et $\Delta k_y$ au plan de reconstruction qui ne correspondent pas à ceux de l'antenne microphonique. En ce sens, une antenne de disposition irrégulière peut être utilisée, bien que cette idée n'est pas discutée en profondeur dans l'article. Un processus de régularisation doit être utilisé lors de la rétro-propagation : dans l'article, la régularisation de Tikhonov est utilisée.

Bien que cette méthode permette une plus grande latitude dans la configuration géométrique du problème, l'utilisation d'une fonction de Green exprimée dans le domaine des nombres d'onde suppose la périodicité spatiale de la convolution et mène à des problèmes de fuites spectrales.

### 2.10.4 Transient Near-field Acoustical Holography (TNAH)

Le *Transient Near-field Acoustical Holography* est une technique introduite par Blais et Ross [8] inspirée des travaux de Hald [32] et de Wu [39]. Son principe repose sur la représentation du champ sonore dans le domaine temporel comme dans le cas du *TDH*. Toutefois, au lieu de résoudre la convolution avec une transformée de Fourier discrète, celle-ci est résolue avec la transformée de Laplace. Une version équivalente à la fonction de Green $g(k_x, k_y, \omega)$ est déterminée analytiquement dans le domaine de Laplace. La mesure est soumise à une transformée de Laplace numérique (*NLT*), qui est équivalente à une GDFT ayant un grand angle $\theta$ imaginaire : les répliques sont alors multipliées par un facteur s'approchant de zéro.

Il est démontré que le formalisme de Laplace permet de réduire les erreurs de fuites spectrales et de recouvrement [40]. Cependant, puisque la fonction de Green est échantillonnée dans le domaine de Laplace, c'est erreurs ne sont pas entièrement supprimée. De plus, étant donné que la procédure s'appliquer seulement sur le domaine temporel, les fuites spectrales et les erreurs de repliement

causées dans le domaine spatial sont toujours présentes étant donnée l'utilisation de la DFT dans ces domaines.

Le TNAH a été appliqué à la résolution du problème direct. Son application à la résolution du problème inverse est plus délicate, puisque le résultat de la déconvolution doit être multiplié par une fenêtre exponentielle croissante : l'erreur de déconvolution devrait alors être amplifiée significativement.

### 2.10.5 **Helmholtz equation least-square method (HELS)**

Une autre méthode d'holographie non-stationnaire basée sur la reconstitution de la source est proposée par Wu [39]. Cette méthode suppose une géométrie en coordonnées sphériques, mais pourrait être adaptée à d'autres géométries.

Le HELS suppose que le champ sonore d'une source sphérique peut être reconstitué par une somme de fonctions harmoniques sphériques. Plus le nombre de termes de la somme est grand, plus la reconstitution de la source est précise. Une méthode des moindres carrés est utilisée pour déterminer l'amplitude des harmoniques sphériques de manière à ce que la propagation du champ sonore déterminé à la source corresponde au champ mesuré. Le nombre de microphones requis dépend du nombre de fonctions harmoniques sphériques considérées. La propagation s'effectue dans le domaine de Laplace et la transformée inverse est calculée au moyen du théorème des résidus.

Une limitation de la méthode est que lorsque la source est de nature complexe, il est difficile de bien représenter le champ sonore avec un petit nombre de termes. Il est intéressant de constater qu'en coordonnées sphériques, étant donnée la symétrie circulaire des variables angulaires, la périodicité spatiale est désirable. Finalement, les erreurs inhérentes à la méthode de convolution utilisée restent présentes.

### 2.10.6 **Méthode des sources équivalentes**

L'idée de la méthode des sources équivalentes est de supposer une distribution de sources ponctuelles sur le plan de reconstruction, puis de résoudre le problème au sens des moindres carrés de telle sorte que le rayonnement des sources concorde avec la condition frontière au plan de

mesure. Un avantage de la méthode est que la distribution des sources ponctuelles sur le plan de reconstruction est arbitraire et ne dépend pas de la distribution des points de mesure.

Zhang et al. ont appliqué cette méthode à la reconstruction de champs acoustiques transitoires [41]. Cependant, la transformation physique considérée dans leur méthode est un simple déphasage suivant d'une normalisation : $\delta\left(t-\frac{R}{c}\right)/R$. Cette transformation n'est qu'une approximation de la véritable fonction de Green exprimée dans le domaine temps-espace à l'équation (2.7). L'applicabilité de la méthode a été démontrée pour des distances de rétro-propagation de l'ordre du centimètre.

Bi et al. ont appliqué une procédure similaire pour déterminer le champ de vitesse particulaire au plan de la source [42]. Ils ont utilisé une interpolation cubique pour résoudre l'équation d'Euler afin de retrouver le champ de pression.

## 2.11 Bilan de la revue de la littérature

En faisant un bilan des différents problèmes de traitement du signal pour la propagation sonore de sources non-stationnaires, une observation importante est que la génération de *wrap-around* est peu considérée malgré que ce soit une cause principale de l'écart avec les valeurs théoriques, particulièrement dans le cas de sources causales. En effet, la majorité des formulations présentées ne considèrent pas le *wrap-around* comme un aspect prioritaire de leur analyse et celui-ci est peu abordé dans la littérature. Il est possible que cette omission soit due à une méconnaissance du problème de la part de certains auteurs. En effet, la terminologie et la différence entre les convolutions linéaires et circulaires ainsi que leurs définitions ne sont pas mentionnées dans la littérature sur l'holographie acoustique.

Une formulation permettant de supprimer entièrement le *wrap-around* permettrait de produire des résultats considérablement plus précis que ceux issus des méthodes qui ne l'éliminent pas. Les méthodes telles le TNAH ou celles qui utilisent le propagateur temps-nombre d'onde permettent de supprimer ou de réduire significativement le *wrap-around* temporel, cependant, aucune ne le supprime également dans le domaine spatial.

Le problème de fuites spectrales est plus amplement discuté dans la littérature. Plusieurs auteurs proposent d'utiliser des filtres et des fenêtres pour réduire l'impact de ces erreurs, malgré que ces techniques ne suppriment pas entièrement les fuites et affectent également les composantes de la

solution recherchée. Il pourrait être plus judicieux d'éviter d'échantillonner les fonctions convoluées dans un domaine autre que celui dans lequel la solution recherchée est représentée. De cette manière, ces erreurs dues à l'application de la transformée discrète seraient évitées. Ceci est une possibilité dans le cas d'holographie temporelle, puisque le champ propagé est représenté selon le temps et que la pression est bien sûr mesurée dans le domaine temps-espace.

Une confirmation de l'importance de ces erreurs de traitement du signal avec les formulations actuelles est que les résultats sont présentés dans une grande majorité des cas que pour de très courtes distances de propagation, généralement de l'ordre de quelques multiples de la discrétisation spatiale. On peut supposer que les erreurs obtenues sont trop significatives pour de plus grandes valeurs de $d$. En effet, les erreurs telles que le *wrap-around* gagnent en importance lorsque la distance de propagation devient plus grande.

La régularisation du problème inverse en NAH a été largement étudiée dans la dernière décennie. La méthode standard est la régularisation de Tikhonov. Plusieurs méthodes ont été appliquées pour déterminer le paramètre de régularisation. Selon la littérature, la méthode bayésienne semble être celle qui donne les meilleurs résultats [30, 37]. Cependant, les performances de ces méthodes dépendent de l'opérateur et du champ mesuré. En établissant une nouvelle formulation du NAH, il devient nécessaire de refaire l'analyse des performances de ces méthodes. De plus, la régularisation parcimonieuse n'a été que très récemment appliquée au problème de NAH stationnaire [43, 44]. Cette méthode n'a pas encore été appliquée à la régularisation de sources non-stationnaires.

Un avantage de plusieurs formulations plus récentes de l'holographie acoustique est que celles-ci permettent de considérer des géométries qui n'impliquent pas une concordance entre le maillage au plan de mesure et au plan de reconstruction. Ceci est avantageux puisque le matériel de mesure utilisé est moins contraint, ce qui peut permettre une économie de coûts. Cependant, ces formulations sont généralement plus coûteuses en calculs : elles impliquent souvent de résoudre des problèmes de moindres carrés sur un très grand nombre de valeurs et elles ne sont pas toujours compatibles avec l'utilisation de la transformée de Fourier rapide (FFT). Ceci est d'autant plus problématique dans la résolution du problème inverse puisque la détermination du paramètre de régularisation est souvent elle-même un problème itératif qui nécessite la résolution du problème inverse à chaque itération. Également, les problèmes non-stationnaires transitoires ont souvent une enveloppe spectrale plus large, qui nécessite une fréquence d'acquisition plus élevée et donc, un

nombre de données plus élevé. Dans ce cas, l'utilisation d'un calcul se basant sur la FFT est fortement suggérée.

# CHAPITRE 3 OBJECTIFS ET MÉTHODOLOGIE

La revue de la littérature démontre que les méthodes actuellement disponibles ne permettent pas d'appliquer l'holographie en champ proche pour de signaux non-stationnaires sans obtenir des écarts avec les prédictions théoriques, et ce, malgré que le problème direct ne soit pas mal posé. Ces erreurs sont dues à la nature discrète et finie des signaux qui engendre du recouvrement et des fuites spectrales. Dans le cadre de ce projet, il est proposé de développer une méthode d'holographie acoustique en champ proche s'appliquant aux signaux non-stationnaires causaux. Cette méthode permettrait de supprimer complètement ces erreurs de recouvrement et de fuites spectrales. Une fois qu'un formalisme adéquat est développé pour le problème direct, les performances des principales méthodes de régularisation et de prédiction du paramètre de régularisation sont déterminées pour la résolution du problème inverse.

## 3.1 Objectif général

Cette thèse vise à développer une formulation d'holographie acoustique en champ proche pour des sources non-stationnaires et causales qui permet de déterminer avec précision le champ sonore.

## 3.2 Objectifs spécifiques

Pour atteindre l'objectif général, les objectifs spécifiques sont définis de la manière suivante :

i. Effectuer les convolutions sans générer d'erreur de recouvrement.

ii. Éviter la génération de fuites spectrales due à l'échantillonnage des signaux convolués dans le domaine de Fourier.

iii. Déterminer les performances et les limites des méthodes de régularisation $\ell^2$ et $\ell^1$ lorsqu'appliquée à la formulation développée en fonction du bruit de mesure et la distance de rétro-propagation.

iv. Évaluer les performances des méthodes de prédiction du paramètre de régularisation lorsque la formulation développée est appliquée aux méthodes de régularisation $\ell^2$ et $\ell^1$.

v. Conserver une efficacité de calcul comparable à celle des méthodes se basant sur l'utilisation de la transformée de Fourier rapide.

vi. Valider la méthode à partir de signaux simulés et de signaux expérimentaux obtenus dans un environnement anéchoïque.

## 3.3 Méthodologie

Dans cette section, les différentes parties du projet de recherche permettant l'atteinte des objectifs sont présentées selon trois axes principaux. En premier lieu, un aperçu des travaux théoriques est exposé. Par la suite, le processus de validation de la méthode théorique est abordé. Il est suivi d'une présentation du montage expérimental et de ses principales caractéristiques.

### 3.3.1 Travaux théoriques

Une formulation du *NAH* non-stationnaire est développée. Cette formulation repose sur l'utilisation de convolutions linéaire 3-D pour calculer la propagation et la rétro-propagation et ainsi déterminer le rayonnement sonore sans générer de recouvrement, et ce, pour chacune des trois dimensions.

Afin de conserver une efficacité de calcul, la convolution linéaire 3-D est déterminée avec l'algorithme de transformée de Fourier rapide. Trois méthodes de calcul de la convolution linéaire sont développées dans cette thèse :

1) La concaténation de séquences de valeurs nulles (*zero-padding*),
2) L'application de la transformée de Laplace numérique,
3) L'application de la transformée de Fourier généralisée.

Les deux dernières méthodes sont abordées dans les deux articles présentés en annexe. Celles-ci sont plus efficaces que la première, mais seulement applicables au problème direct.

Ces algorithmes de convolution nécessitent que les deux fonctions convoluées soient définies dans le domaine temps-espace, c'est-à-dire dans le domaine dans lequel la solution est représentée. Ainsi, en échantillonnant la fonction de Green dans ce domaine, on évite la génération d'erreurs de fuites spectrales. Le formalisme derrière la méthode proposée est donc développé en ce sens.

Finalement, les méthodes de régularisation $\ell^2$ et $\ell^1$ sont appliquées. Pour la détermination du paramètre de régularisation, la courbe en L, la validation croisée généralisée ainsi que l'approche Bayésienne sont considérées pour la régularisation de Tikhonov, et la méthode de la courbe en L est considérée pour la régularisation $\ell^1$.

Le principal outil de calcul utilisé pour l'ensemble des travaux est MATLAB.

### 3.3.2 Validations

La validation de la méthode temps-espace se fera de deux manières : par simulations numériques et expérimentalement. Dans le cas des simulations, la moyenne quadratique (*Root mean square*-RMS) de la différence entre un signal de référence et le signal obtenu avec chacune des méthodes évaluées est considérée. Puisque les signaux comparés sont tridimensionnels, on somme l'écart sur les trois dimensions pour obtenir l'erreur RMS. Le signal de référence est généré avec un modèle de piston dans un mouvement transitoire [45]; ce signal est exact dans le champ proche et lointain de la source. On génère ainsi le champ sur divers plans avec le modèle théorique, qu'on utilise pour comparer les résultats en propagation et en rétro-propagation. On considère un profil de déplacement sinusoïdal amorti pour le piston, de manière à se rapprocher du cas de la plaque expérimentale.

Dans le cas des signaux expérimentaux, la validation n'est pas quantifiée. Elle est effectuée en observant le champ calculé en propagation ou en rétro-propagation, et en y identifiant la présence d'artefacts de traitement du signal. Par exemple, la causalité de la solution obtenue est constatée en observant la présence de répliques indésirables dans les signaux temporels ou dans les reconstructions spatiales.

### 3.3.3 Mesures expérimentales

Le montage expérimental est constitué d'un banc d'essai composé d'une glissière sur laquelle se rattachent 64 microphones, d'un impacteur, d'un microphone de référence, d'une plaque de poly(méthacrylate de méthyle) et d'un système d'acquisition. Les microphones sont espacés d'un centimètre centre à centre. L'impacteur consiste en une tige métallique à embout arrondi qui est maintenu en suspension au-dessus de la plaque par un électro-aimant. La tige métallique est introduite dans un tube qui guide sa trajectoire. L'impacteur permet une bonne répétabilité des signaux générés : sur vingt mesures, on obtient moins de 2% d'écart maximal sur l'ensemble des valeurs au microphone de référence.

La glissière permet d'effectuer une translation des microphones de manière à enregistrer le son sur le plan complet. Pour chaque position, un enregistrement est effectué et le microphone de référence permet de synchroniser les acquisitions les unes par rapport aux autres.

Les prises de mesures sont effectuées dans un environnement anéchoïque pour minimiser le bruit de fond et simuler un champ libre. L'antenne de microphone couvre une superficie au moins deux fois plus grande que celle de la plaque de manière à minimiser les erreurs de troncature.

Les 65 microphones utilisés sont certifiés « classe 1 » et sont jugés de précision suffisante, soit d'environ 0.5 dB sur la plage fréquentielle d'intérêt. La plage dynamique couvre les amplitudes sonores appropriées. Le système d'acquisition est un PXIe-1082 de *National Instruments* qui permet une bande passante de 8 G/s et des fréquences d'acquisitions au-delà du 50 kHz requis.

Un impact est généré par la tige métallique frappant la plaque de plexiglass. Le mouvement de la plaque est amorti en une dizaine de millisecondes. Un exemple de signal impulsionnel est présenté à la Figure 3.1.

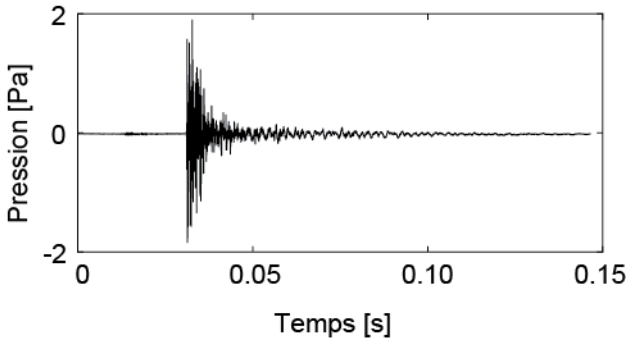


*Figure 3.1 : Exemple de variation de la pression en fonction du temps obtenu expérimentalement*

L'amplitude est initialement nulle, puis augmente soudainement pour s'amortir et devient relativement faible environ 10 millisecondes après l'impact.

# CHAPITRE 4 ARTICLE 1 : TIME DOMAIN NEARFIELD ACOUSTICAL HOLOGRAPHY WITHOUT WRAP-AROUND ERROR AND SPECTRAL LEAKAGE FOR FORWARD PROPAGATION

J.-M. Attendu, A. Ross, publié dans le *Journal of the Acoustical Society of America*, 141 (2), pp. 1039-1050, 2017.

## 4.1 Préface

Cet article établit les principales bases de cette thèse, puisque la formulation qui repose sur l'utilisation de la convolution linéaire 3-D y est développée. Cette formulation permet d'atteindre les objectifs *i* et *ii*, car elle supprime complètement les erreurs de recouvrement et de fuites spectrales. Cela est réalisé tout en conservant une efficacité de calcul, ce qui répond à l'objectif *v*. On y démontre que le problème direct peut être résolu sans générer d'erreurs significatives, et ce, peu importe la distance de propagation considérée. On compare les résultats avec ceux de d'autres formulations pour le problème direct : les très faibles erreurs obtenues mettent en évidence que les problèmes de traitement du signal présents dans les autres formulations sont évités avec la formulation proposée. On y établit certaines balises concernant la taille de la fonction de Green selon la distance de propagation, de manière à éviter les erreurs de troncature. Finalement, la méthode proposée est appliquée à un cas expérimental. Ces validations par simulations et mesures expérimentales permettent d'atteindre l'objectif *vi* pour le problème direct.

## 4.2 Abstract

Nearfield acoustical holography can be used to visualize time domain evolution of non-stationary sound fields. It requires sampling of the pressure field with a microphone array to calculate the sound field on other parallel surfaces, either toward the source or farther away from it. This paper focuses on forward propagation of fields sampled with a planar array. Several formulations exist to perform such calculation; however, despite being mathematically well-posed, they generate significant errors, especially for larger propagation distances. These errors are mainly due to wrap-around and spectral leakage. It is shown in this paper that generation of leakage can be avoided by sampling the analytical Green's function in the time and space domains. Such formulation allows the application of linear convolution, which completely suppresses wrap-around errors. The results

obtained from the proposed formulation are compared to the analytical solution of a baffled piston with transient displacement. The average discrepancy is below 3.5 % for propagation distances of 5 to 55 cm. In comparison, the relative error obtained with other propagators varies from 6 % to 73 %, depending on the propagation distance and method used. Finally, the proposed formulation is applied to experimental data obtained from an impacted plate.

## 4.3 Introduction

Non-stationary acoustic sound fields can take many forms: impact noises, engine run-ups, squeal noises, musical sounds, machinery-induced noises, etc. Their abundance in industrial environments, transportation, construction, and their impact on health and safety are some of many justifications to the relevance of monitoring these noises. Visualization of sound fields can be achieved with nearfield acoustical holography (NAH) [3] [2], by acquiring the pressure field with an array of microphones positioned in the nearfield of the source. The sampled pressure field acts as a boundary condition to solve the wave equation using Green's formulation. In planar NAH, the microphones are distributed uniformly on a rectangular plane, and the sound field is determined on other collinear planes toward (inverse problem) or away from the source (direct problem).

The standard NAH formulation supposes a stationary acoustic source. Consequently, the time domain circular convolution implicitly used in the Green's formulation assumes periodic signals. Because the statistical properties of stationary signals do not fluctuate significantly in time, the propagated sound field is represented as a function of the frequency $\omega$, at different spatial coordinates $x$ and $y$. However, this is not applicable for non-stationary signals as their spectral components vary in time. Consequently, non-stationary sound fields must be represented in the time domain.

Such idea was proposed by Hald, who considered the application of an inverse discrete Fourier transform (DFT) over $\omega$ to the standard NAH formulation to obtain time domain representation [32]. This method is known as Time-domain acoustical holography (TD-NAH). However, TD-NAH used with the standard NAH propagator is subject to important signal processing errors, as noted by the author.

TD-NAH was further investigated by De La Rochefoucauld, Melon and Garcia [7], using four different Green's function formulations defined in combination of time/frequency and space/wave

number domains, including the time and space domain (TSD) Green's function used in the present paper. They performed a study over various parameters such as the propagation distance, characteristics of the source and of the microphone array. The performance of the studied propagators was shown to vary with the parameters studied, but in all cases, the error becomes significant as the propagation distance increases.

Blais and Ross proposed to trade the time domain discrete Fourier transform (DFT) in TD-NAH for the numerical Laplace transform [46]. Because of its mathematical properties, the Laplace transform is well suited for the processing of transients, since it reduces errors due to leakage and attenuates the aliased replicas causing wrap-around [47], however, it does not suppress these errors completely. The numerical Laplace transform was also previously used by Wu for spherical propagation of acoustic transients [39].

Thomas and Grulier [6] [34] introduced another formulation, Real-time nearfield acoustical holography (RT-NAH). It uses a propagator expressed in the time/wave number domain. Such propagator was also used by Forbes, Letcher and Stepanishen [35]. With RT-NAH, the limitation of TD-NAH to short signals is overpassed by performing linear combination of successive time domain convolutions. The authors noted the presence of important spectral leakage errors in their results, and evaluated the performances of different filters to reduce these errors.

Fundamentally different approaches were also proposed for the representation of non-stationary sound fields, including the use of the equivalent sources method [41] [42] [48] [38] or the wavelet transform [49].

The errors inherent to the described TD-NAH methods take root in the mathematical foundations of the problem. In all cases presented in the literature, some conditions regarding the evaluation of the analytical convolution with a discrete and finite formulation are not respected.

The purpose of the present paper is thus:

1. To clearly establish the causes of fundamental signal processing errors present in the existing time domain NAH formulations.
2. To present a methodology using the TSD propagator that suppresses both leakage and wrap-around errors, thus considerably improving the TD-NAH results.
3. To provide a rule ensuring that Green's function truncation errors remain low.

4. To present a validation of the proposed method with both simulated and experimental results.

## 4.4 Time domain NAH and its formulations

The different propagators for time domain NAH are described in this section. Hansen presented the rigorous mathematical derivation for time domain Green's function formulation in [12]. It is based on the use of the Dirichlet Green's function $G$ to calculate the pressure field $p$ on the plane $z = z_0 + d$ from the known pressure field on the parallel plane $z = z_0$. The operation involved in this calculation is a three-dimensional convolution:

$$p(t, x, y, z_0 + d) = \iiint_{-\infty}^{\infty} p(t', x', y', z_0) \cdot G(t - t', x - x', y - y', d) dt' dx' dy'. \qquad (4.1)$$

Euler's relation between sound pressure and particle velocity [13] can be applied to either the Green's functions or the measured pressure field to obtain the particle velocity vector upon performing the convolution. Sound intensity can be obtained from the product of the pressure and the velocity.

Because of the convolution property of the Fourier (or Laplace) transform [5], the analytical convolution is equivalent to a product in the Fourier space. This property can be applied to any combination of the three dimensions $x$, $y$ and $t$, which leads to the various formulations for the Green's propagators presented in Tableau 4.1.

*Tableau 4.1: Green's functions for the different TD-NAH formulations*

| *Green's functions* |
| --- |
| $G(\omega, k_x, k_y, d) = e^{i \cdot d \cdot \sqrt{k^2 - {k_x}^2 - {k_y}^2}}.$ (4.2) |

*Tableau 4.1 (suite): Green's functions for the different TD-NAH formulations*

| | |
|---|---|
| $$G(s,k_x,k_y,d) = e^{-\kappa d}e^{i\cdot d\cdot k_z'},$$ $$k_z' = \left(\frac{a+\sqrt{a^2+b^2}}{2}\right)^{1/2},$$ $$\kappa = \left(\frac{-a+\sqrt{a^2+b^2}}{2}\right)^{1/2},$$ $$a = \frac{\omega^2-\sigma^2}{c^2} \text{ and } b = -\frac{2\omega\sigma}{c^2},$$ $$s = \sigma + i\omega.$$ | (4.3) |
| $$G(t,k_x,k_y,d) =$$ $$\delta(t-\tau) - \tau\Omega_r^2\frac{J_1\left(\Omega_r\sqrt{t^2-\tau^2}\right)}{\Omega_r\sqrt{t^2-\tau^2}}\Gamma(t-\tau),$$ $$\tau = d/c,$$ $$\Omega_r = c\sqrt{k_x^2+k_y^2}.$$ | (4.4) |
| $$G(t,x,y,d) = \frac{d}{2\pi c}\left(\frac{c}{R^3}+\frac{1}{R^2}\frac{\partial}{\partial t}\right)\delta(t-R/c).$$ $$R = \sqrt{x^2+y^2+d^2}$$ | (4.5) |

Eq. (4.2) is the standard NAH propagator proposed by Williams and Maynard [3], with $d$ being the propagation distance, $k$ expresses the wave-number and $k_x$, $k_y$ are its $x$ and $y$ components. Eq. (4.3) is the propagator proposed by Blais and Ross [46], where $s$ is the time variable expressed in the Laplace domain, $\sigma$ is a positive damping constant, $c$ is the sound velocity and $\omega$ is the angular frequency. The RT-NAH propagator proposed by Grulier and Thomas [6] is expressed in (4.4), where $J_1$ corresponds to the Bessel function of the first kind and $\Gamma(t)$ is the Heaviside function.

The TSD propagator derived by Hansen [12] and applied to NAH by La Rochefoucauld [7] is presented in Eq. (4.5).

The Green's functions expressed in Tableau 4.1 are all equivalent and lead to the same results when the convolution is calculated analytically. However, this does not hold true with sampled signals. In the next section, we explain the cause of these discrepancies, which are due to signal processing errors caused by the discrete and finite nature of the convolved signals.

## 4.5 Signal processing errors in TD-NAH

The main causes of signal processing errors in TD-NAH are wrap-around, spectral leakage and truncation errors. In this section, the first two phenomena are described in the scope of TD-NAH and it is explained how the use of the TSD propagator can suppress these errors. Time and spatial domains truncation errors are also addressed in this section, since it becomes the main source of error with the TSD formulation. It is assumed that the sampling of the studied pressure field respects the Nyquist criterion in both the time and spatial domains. Such assumption is reasonable when the mechanical sources studied are bandlimited.

### 4.5.1 Finite convolution and wrap-around error

The key operation performed in the calculation of the propagated sound field is the numerical convolution expressed in (4.1). A problem arises as the analytical convolution supposes integration over $\mathbb{R}$, but the sampled signals considered in the numerical convolution are measured over finite domains.

#### 4.5.1.1 Linear and circular convolution

Two classical assumptions can be considered to overcome this limitation; either the finite signals are supposed null outside of their domain of definition, or they are repeated periodically. These assumptions lead to two different numerical convolution results, respectively linear and circular convolution [22]. Wrap-around error is generated when circular convolution is applied, even when the physical conditions of the problem correspond to the assumptions of the linear convolution, that is when the actual signals are supposed null outside of the sampling period. Here, the discrete functions $f(n)$ and $g(n)$ defined over $\{0, \cdots, N-1\}$ are considered, and it is supposed that their values are null outside of this interval.

The mathematical expression for linear convolution is:

$$h_{lin}(n) = \sum_{m=0}^{N-1} f(m)g(n-m), \tag{4.6}$$

with $n \in \{0, \dots, 2N-2\}$ and $g(n-m) = 0$ for $n-m < 0$ and $n-m > N-2$. The linear convolution of two sequences of $N$ samples results in a sequence of $2N-1$ samples.

Circular convolution is expressed as:

$$h_{circ}(n) = \sum_{m=0}^{N-1} f(m)g((n-m)_{mod\,N}), \tag{4.7}$$

with $n \in \{0, \dots, N-1\}$. Because of an inherent property of the DFT [5], the circular convolution can also be obtained by taking the DFT (or the FFT) of both signals and by multiplying them in the Fourier domain:

$$h_{circ}(n) = DFT^{-1}\{DFT\{f_n\} \cdot DFT\{g_n\}\}. \tag{4.8}$$

The calculation of the convolution with the FFT has the advantage that the computation is considerably more efficient [17], especially considering the three-dimensionality and the large amount of data computed in NAH.

#### 4.5.1.2 Wrap-around in NAH

In planar NAH, considering a source centered with the measurement aperture at $x = y = 0$, the amplitude of the radiated sound field is attenuated as $x, y \to \pm\infty$. Since it is supposed that the microphone array is large enough to capture (most of) the radiated sound field, we therefore assume that the signal is null outside of its domain of definition and thus, that the linear convolution assumption is representative of the spatial convolution.

For the time domain convolution, if the pressure field is causal and quickly decreases to zero, such as with impact noises, linear convolution assumption is also adequate. For longer non-stationary signals, block convolution is applied: it consists in recombining a series of successive linear convolution of short sequences to obtain the linear convolution of a longer sequence. This is possible because convolution is a linear operation. The outcome of using this method is that the

time domain representation is no longer limited by the length of the sample. Thomas proposed a similar approach for dealing with long non-stationary signals [6]. The two standard block convolution methods are the overlap-add and overlap-save methods [22].

To illustrate the effect wrap-around can have on the propagated field, time domain NAH using circular convolution is applied to a short causal impulse signal generated using the baffled piston model described in Section 4.7. The standard NAH propagator expressed in (4.2) is used and the circular convolution is performed by multiplying the signals in the Fourier domain. The results for a propagation distance of $d = 0.30$ m are presented in Figure 4.1, along with a comparative analytical solution.

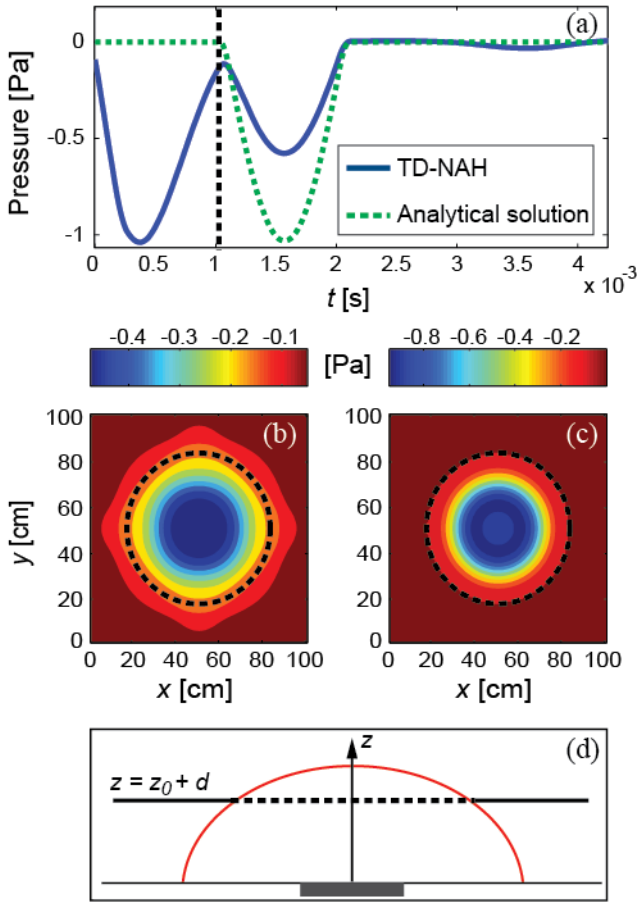


*Figure 4.1 : Illustration of wrap-around error in the case of a causal signal. a) In the time domain on the central axis. b) In the spatial domain with the standard NAH method ($t = 1.8$ ms*

*and $d = 0.30$ m). c) Analytical results in the spatial domain ($t = 1.8$ ms and $d = 0.30$ m). d) Projection of a propagating wave front over the x-z plane passing through the center of the piston*

Wrap-around error is particularly problematic in this case since the studied signal is causal and consequently, the amplitude of the propagated field should be null before the wave front reaches the propagation plane. This event is represented in the time domain in Figure 4.1a) by the black dotted line at $t_c \approx 1$ ms. Any variation of the sound pressure prior to $t_c$ does not comply with the causality of the system; however, an aliased replica created by wrap-around is clearly present.

In the spatial domains, the black dotted circles of Figure 4.1b) and Figure 4.1c) delimit the physical propagation of the wave front at the represented time $t = 1.8$ ms. Again, wrap-around error produces non-zero components beyond this limit in Figure 4.1b). A projection of the propagating wave front over the $z$ axis is schematized in Figure 4.1d). The baffled piston is illustrated in gray at $z = 0$. The propagating wave front is shown in red for an arbitrary time $t$. The wave front crosses the observation plane $z = z_0 + d$ and the dotted line delimits the boundaries of the non-zero pressure field on the plane.

#### 4.5.1.3 Implementation of linear convolution

A simple way of implementing 3-D linear convolution is to double the values of both the pressure field and the Green's function with trailing zeros with respect to each dimension. Convolution is then calculated by means of applying three-dimensional fast Fourier transform (FFT) to both signals before multiplying them in the Fourier domain. Finally, TSD representation is obtained by applying the inverse FFT. This zero-padding method is mathematically equivalent to calculating the convolution with the summation form in (4.6).

It is important to understand that if only the pressure field is zero-padded, as suggested by some authors, wrap-around error is not suppressed (although it can be reduced). To support this, the initial pressure field used for Figure 4.1 was zero-padded prior to the convolution; however, this did not prevent the generation of considerable wrap-around errors.

By expressing the Green's function in the time and space domains (i.e. eq. (4.5)), it is possible to apply zero-padding to the Green's function prior to the convolution. On the contrary,

multiplication with the Green's function expressed in one or more frequency domains supposes its periodicity and circular convolution result is obtained, causing wrap-around error.

The zero-padding method can also be replaced by other more computationally efficient means of obtaining multi-dimensional linear convolution, in particular by using the generalized discrete Fourier transform [50] or the numerical Laplace transform [47]. The computational gain of those methods is around 50 %.

### 4.5.2 Spectral leakage

Another important recurring problem in NAH is spectral leakage. It is due to the discretization and truncation of the analytic Green's function expressed in the Fourier domain; the application of the inverse discrete Fourier transform generates errors. In this section, the presence of spectral leakage in NAH is highlighted by a simple derivation. The derivation is done with respect to the time and frequency, but similar errors are obtained from the inverse spatial DFTs. A representation of Green's functions in the time and space domains illustrates the presence of leakage. It is explained how the use of TSD-NAH suppresses these errors. The analytical Fourier transform pair considered in the derivation is defined by Eq. (4.9) and Eq. (4.10) for an arbitrary function:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) \cdot e^{-i\omega t} dt, \tag{4.9}$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \cdot e^{i\omega t} d\omega. \tag{4.10}$$

#### 4.5.2.1 Derivation

It is assumed that the measured sound field expressed in the Fourier domain $P(\omega, z_0)$ is known. To obtain the propagated sound field, $P(\omega, z_0)$ is multiplied with the Green's function $G(\omega)$. In practice, these functions are discrete, so they are multiplied with a Dirac comb $\Pi(\omega)$:

$$P(m, z_0 + d) = \Pi(\omega) \cdot P(\omega, z_0) \cdot G(\omega, d), \tag{4.11}$$

where $m \in \mathbb{Z}$, and

$$\Pi(\omega) = \sum_{m\in\mathbb{Z}} \delta(\omega - m\Delta\omega). \tag{4.12}$$

The inverse analytical Fourier transform of (4.12) is another Dirac comb:

$$\Pi(t) = \mathcal{F}^{-1}\{\Pi(\omega)\} = \frac{1}{\Delta\omega}\sum_{n\in\mathbb{Z}} \delta\left(t - \frac{2\pi n}{\Delta\omega}\right). \tag{4.13}$$

The relationship between the discretization in the Fourier and space-time domains is obtained by considering the convolution property of the Fourier transform: the multiplication with the Dirac comb in the Fourier domain corresponds to a convolution with (4.13) in the time domain:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} P(\omega, z_0 + d)\Pi(\omega)\cdot e^{i\omega t} d\omega = \int_{-\infty}^{\infty} p(t - t', z_0 + d)\Pi(t')dt'. \tag{4.14}$$

Equation (4.14) reduces to:

$$\frac{\Delta\omega}{2\pi}\sum_{m\in\mathbb{Z}} P(m\Delta\omega)\, e^{im\Delta\omega t} = \sum_{n\in\mathbb{Z}} p\left(t - \frac{2\pi n}{\Delta\omega}\right). \tag{4.15}$$

By isolating $p(t)$ by means of taking $n = 0$ in (4.15), we obtain:

$$p(t) = DFT^{-1}\{P(m\Delta\omega)\} + E_1 + E_2, \tag{4.16}$$

with,

$$E_1 = \frac{\Delta\omega}{2\pi}\left(\sum_{m=-\infty}^{\frac{N}{2}-1} P(m\Delta\omega)\, e^{im\Delta\omega t} + \sum_{m=N/2}^{\infty} P(m\Delta\omega)\, e^{im\Delta\omega t}\right), \tag{4.17}$$

and

$$E_2 = -\sum_{n\in\mathbb{Z}^*} p\left(t - \frac{2\pi n}{\Delta\omega}\right). \tag{4.18}$$

The first term of the right-hand side of (4.16) is the inverse discrete Fourier transform of $P(m\Delta\omega)$. $E_1$ and $E_2$ represent the errors resulting from the Fourier domain truncation and discretization, respectively.

In NAH, the usual means of reducing these errors is by filtering the pressure field, or by multiplying it with a window prior to the application of the DFT. For example, Thomas and his co-authors [34] found that the use of Chebyshev and Kaiser filters improves the representation in the time domain. Also, Scholte et al. found that extrapolating the pressure field by means of zero-padding the spatial domain improves the reduction of leakage due to the multiplication with a Tukey window [51]. These methods reduce leakage by filtering high frequency errors, but at the cost of altering the spectral content of the solution. Furthermore, it can be difficult to find the ideal filter or window, and such treatments do not suppress $E_1$ and $E_2$ completely.

By convolving the sampled pressure field with Green's function expressed directly in the time and space domain, errors due to the discretization and truncation of the Green's function in the frequency domain are avoided, which greatly improves the TD-NAH results.

#### 4.5.2.2 Illustration of leakage in the Green's function

To illustrate the impact of spectral leakage in the Green's functions, inverse DFT is applied to the Fourier domain Green's function expressed in (4.2). Its spatial representation at time $\tau = 2$ ms is compared to the time and space Green's function (eq. (4.5)). A propagation of $d = 0.15$ m over the $z$ axis is considered.

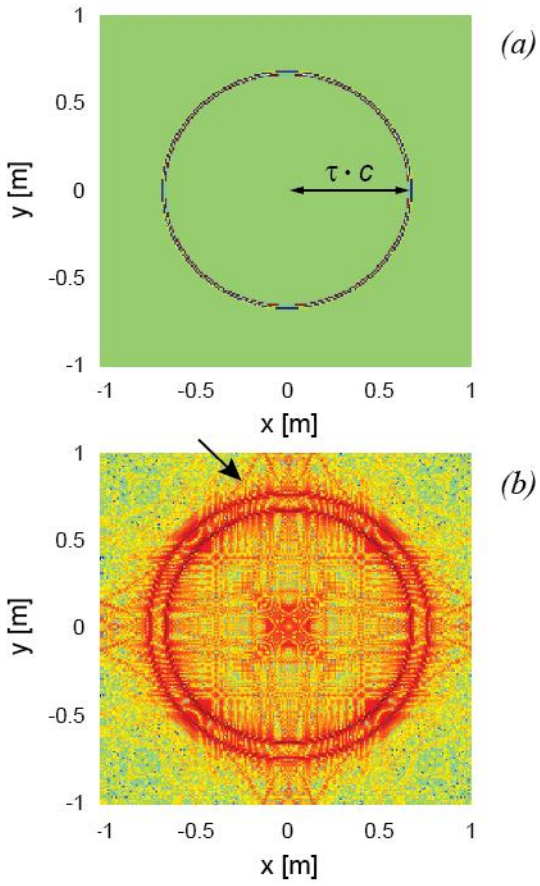


*Figure 4.2 : Comparison of Green's functions expressed in the spatial domain at $t = 2$ ms. a) TSD-NAH formulation. b) Obtained from the inverse DFT of the analytical expression in the Fourier domain (logarithmic scale)*

The TSD-NAH propagator is illustrated in Figure 4.2a). The Dirac delta functions distributed radially produce a spatial shift of the measured pressure field upon convolution. This shift corresponds to the propagation distance ($\tau \cdot c$) at the studied time sample $\tau$. At other positions, the amplitude is null, and consequently, no causal error is produced by the convolution.

In Figure 4.2b), the inverse DFT formulation is presented in logarithmic scale to highlight the presence of the errors. Most of the amplitude is distributed radially (inner full circle) with respect to the center, which corresponds to the spatial shift presented in Figure 4.2a). Important ripples due to the truncation error $E_1$ are present over the whole domain, and especially near the center. The aliased replicas due to the Fourier domain discretization error $E_2$ are also visible in Figure 4.2b), for example, as denoted by the black arrow. Aliased replicas in the time domain add

a time-shifted duplication of the Green's function in the $x$ and $y$ representation, which corresponds to the outer circle in Figure 4.2. The radius of the outer circle varies with the time domain sampling period.

It is clear that convolution with such propagator would result in important discrepancies with the analytical solution. By sampling the analytical Green's function directly in the domains in which the propagated sound field is to be represented, that is the time and space domains in this case, leakage errors are avoided.

### 4.5.3 Truncation errors from the time and spatial domains

The linear convolution supposes that the convolved signals are null outside of their finite domains. In practice, this assumption is not respected, as some (small) proportion of the pressure field is not measured by the microphone array in both the time and the spatial domains. However, convolution of the truncated parts with the Green's function also contributes to the propagated sound field, as illustrated in Figure 4.3 for the $x$ domain. This contribution is not considered in the NAH calculation, and is the cause of truncation errors.

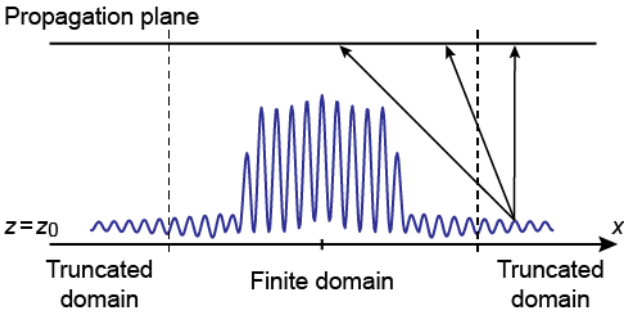


*Figure 4.3 : Illustration of the contribution of the truncated pressure field to the propagation plane*

In Figure 4.3, the vertical dashed lines delimit the finite NAH domain on both the measurement plane ($z = z_0$) and the propagation plane of an arbitrary source. Point sources in the truncated zones contribute to the propagated field, as illustrated by the arrows, however, these contributions decrease with the propagation distance. For this reason, the truncation errors are more important near the boundaries of the finite domain on the propagation plane. In addition, the truncated data

closest to the boundaries are susceptible to cause more significant errors. Likewise, the truncation of the Green's function also generates errors, as discussed in Section IV-B.

## 4.6 Implementation of TSD-NAH

In this section, an overview of the main operations involved in TSD-NAH is presented. Then, a formula is derived to determine the minimum size of the Green's function needed to avoid important truncation errors. This is followed by a discussion on the application of the time domain differential operator of (4.5).

### 4.6.1 **Main operations in TSD-NAH**

The key operation in TSD-NAH are presented using the diagram in Figure 4.4. Here, it is supposed that linear convolution is calculated using the zero-padding method.

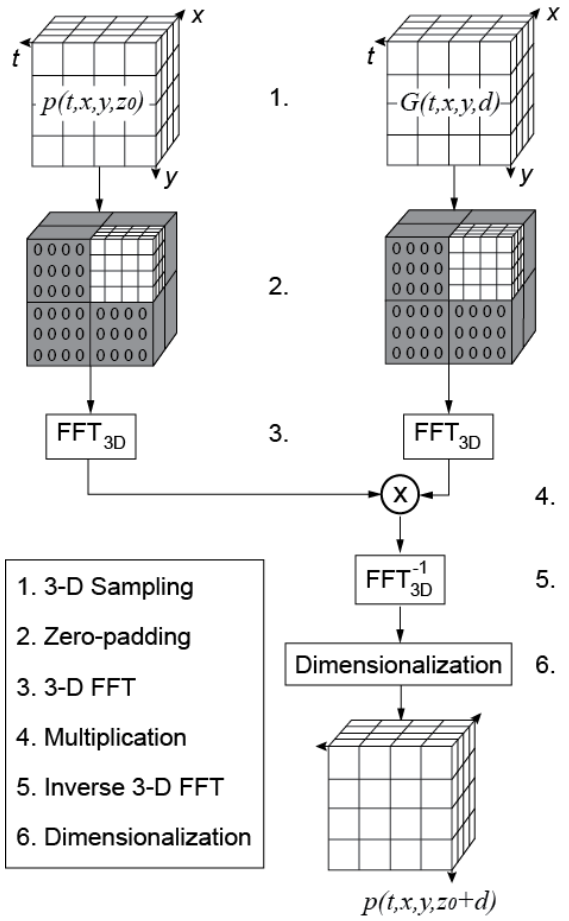


*Figure 4.4 : Diagram of the operations involved in TSD-NAH*

The first step is to sample the pressure field in three-dimensions. This takes the form of a sequence of snapshots of the instantaneous pressures over the microphone array area sampled at frequency $f_s$. The TSD-NAH Green's function is also sampled accordingly. Then, zero-padding is applied to both the Green's function and the pressure field, doubling the size of the signals with respect to each of the three dimensions. Three-dimensional FFT is applied to both signals, which are then multiplied in the Fourier domain. An inverse 3-D FFT is performed to obtain the TSD representation. Finally, the propagated sound field is truncated in order to retain only the samples corresponding to the original time and space domain windows of the measured pressure field. The truncated values can be kept to visualize a wider representation of the sound field, but they could contain truncation errors, as discussed in Section 4.7.

### 4.6.2 **Effect of the propagation distance on the finite spatial domain**

The amplitude of the time and space Green's function expressed in (4.5) decreases as the magnitude of $x$ and $y$ increases. In fact, for high values of $x$ and $y$, the term in $1/R^2$ governs the amplitude of the function. To apply linear convolution, the Green's function is truncated in order to obtain a finite sequence, and it is supposed that the convolved sequences are null outside of their domain of definition. Thus to avoid truncation errors, the finite domain must be wide enough to encompass most of the Green's function amplitude.

It is important to notice that the extent of the Green's function over the spatial domain increases with the propagation distance $d$. The term with $1/R^2$ is illustrated in Figure 4.5 for the spatial domain $[-1,1]$ m with $d = \{0.02, 0.15, 0.35, 0.50\}$ m. Because of the radial symmetry in $x$ and $y$, Figure 4.5 presents the projection over the $x$ axis only.

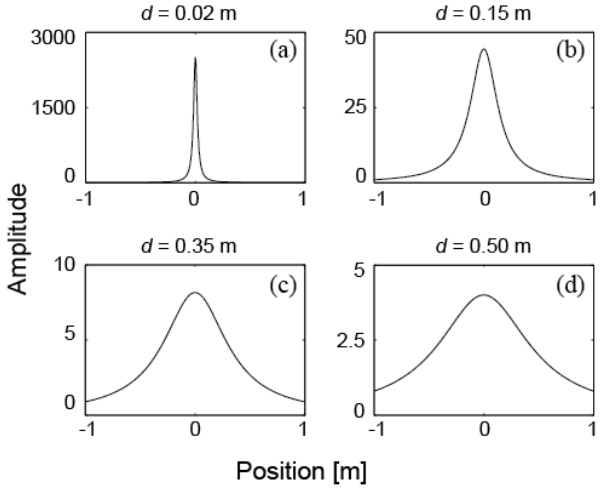


*Figure 4.5 : Spatial distribution of the finite $1/R^2$ function with respect to the propagation distance. a) $d = 0.02\ m$. b) $d = 0.15\ m$. c) $d = 0.35\ m$. d) $d = 0.50\ m$*

It can in fact be observed in Figure 4.5d) that for the largest propagation distance $d = 0.50$ m, an important part of the function is truncated. The contribution of the truncated parts is suppressed from the convolution result, which generates discrepancies with the analytical solution.

A formula is derived to calculate a minimal size of the spatial domains with respect to $d$. The derivation supposes that the term with $1/R^2$ is governing the behavior of the Green's function for higher values of $x$ and $y$, thus neglecting the term with $1/R^3$. Considering the radial symmetry, the governing term projected over $x$ is:

$$\Lambda = \frac{1}{x^2 + d^2}. \tag{4.19}$$

By integrating Λ over ℝ, we obtain:

$$\int_{-\infty}^{\infty} \Lambda \, dx = \frac{\pi}{d}. \tag{4.20}$$

Then, Λ is integrated over the finite domain bounded at $\pm x_{max}$:

$$\int_{-x_{max}}^{x_{max}} \Lambda \, dx = \frac{2}{d} atan\left(\frac{x_{max}}{d}\right), \tag{4.21}$$

where $x_{max}$ is expressed in meter. We define $\alpha$ corresponding to the ratio of the amplitude encompassed in the finite domain over the total amplitude over ℝ. The ratio is obtained by dividing (4.21) by (4.20), giving:

$$\alpha = \frac{2}{\pi} atan\left(\frac{x_{max}}{d}\right). \tag{4.22}$$

The boundaries of the finite domain are obtained by isolating $x_{max}$ in (4.22):

$$x_{max} = d \cdot tan\left(\frac{\alpha\pi}{2}\right) \tag{4.23}$$

We have found that $\alpha \approx 0.85$ is an adequate trade-off between the computational cost due to the size of the finite domain and the truncation error, leading to:

$$x_{max} \geq 4 \cdot d. \tag{4.24}$$

To support this, a normalized $1/R^2$ function, analogous to the term that governs the Green's function in (5), is presented with respect to $x/d$ in Figure 4.6. The dashed vertical lines represent the truncation proposed in (4.24).

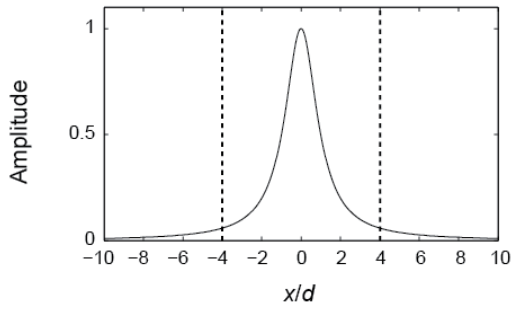


*Figure 4.6 : Illustration of the normalized $1/R^2$ function with respect to $x/d$*

The slope of the function is steeper near the origin and decreases significantly for higher values of $|x|$. We observe in Figure 4.6 that the proposed truncation occurs at a point at which the slope is relatively small and the amplitude is considerably lower than that at the origin. Even though the truncated parts ($|x| > 4$) correspond to about 15 % of the total area under the entire curve, it is fairly well distributed over the whole truncated domain. The amplitude decreases gradually from 0.06 at the truncation to 0 at infinity. The farther the truncated data are from the center, the less impact they have on the convolution results. Consequently, assuming the amplitude of the measured pressure field is spatially localized, the contribution of the entire truncated parts to the convolution integral is not significant. The same trade-off value is used for $y_{max}$.

It is important to understand that the size of the measured pressure field, corresponding to the size of the microphone array, does not have to fit the Green's function; the measured pressure field can simply be zero-padded to reach $x_{max}$. Since the Green's function is known analytically, the consequence of respecting (4.24) only implies an increased computational cost.

### 4.6.3 **Differentiation error**

The time domain derivative in (4.5) is applied to the measured pressure field. We found that the fourth order centered finite difference operator is adequate to obtain the derivative of $p(t)$ without generating significant errors. The formula is given in equation (4.25) [52]:

$$p_0' = \frac{(p_{-2} - 8p_{-1} + 8p_1 - p_2)}{12\Delta t} + O(\Delta t^4). \tag{4.25}$$

It is supposed that the sampling frequency is high enough for the pressure variation to be smooth with respect to $t$. It is usually the case since the sampling frequency of modern acquisition systems is considerably higher than the limited bandwidth of the studied mechanical source. However, in the presence of important numerical noise, a data smoothing algorithm must be applied to avoid amplification of the noise due to the differentiation. The approach proposed by Stickel [53] was previously tested by the authors to successfully suppress differentiation error in TSD-NAH due to white Gaussian noise added numerically. One asset of this approach is that it relies on the use of regularization function that does not assume the periodicity of the signal as usual filtering does.

Following Eq. (4.25), TSD-NAH propagator expressed in (4.5) can be rewritten by applying the differential operator to the Dirac delta. Equation (4.5) thus becomes:

$$G = \frac{d}{2\pi c}\left(\frac{c \cdot \delta(\tau)}{R^3} + \frac{D}{R^2}\right), \tag{4.26}$$

with

$$D = \frac{1}{12\Delta t}\left(\delta(\tau - 2\Delta t) - 8\delta(\tau - \Delta t) + 8\delta(\tau + \Delta t) - \delta(\tau + 2\Delta t)\right) \tag{4.27}$$

and

$$\tau = t - \frac{R}{c}. \tag{4.28}$$

## 4.7 Comparison with Analytical solution to the transient radiation of a baffled piston

In this section, the effectiveness of TSD-NAH is analyzed by comparing the resulting pressure fields with those obtained with the analytical solution of a circular baffled piston sustaining a transient excitation. Propagation results obtained using other Green's functions introduced above are also compared.

### 4.7.1 **Baffled piston radiation model**

The analytical model for transient radiations of a circular baffled piston was developed by Stepanishen [45]. The model is based on Green's function formulation and is exact in both the nearfield and the far-field. It is applicable to any displacement of the piston. Stepanishen showed that the pressure field with respect to time and the position $R$ can be obtained by convolving the acceleration of the motion of the piston with the impulse response $h(R,t)$:

$$p(R,t) = \rho_0 h(R,t) * \frac{\partial^2 x(t)}{\partial t^2}, \tag{4.29}$$

where $\rho_0$ is the density of air. The geometry associated to the baffled piston is presented in Figure 4.7, where $r$ is the projection of $R$ over the $x$ and $y$ plane, $a$ is the radius of the piston and $l$ is the propagation distance over the $z$ axis to the so-called measurement plane.

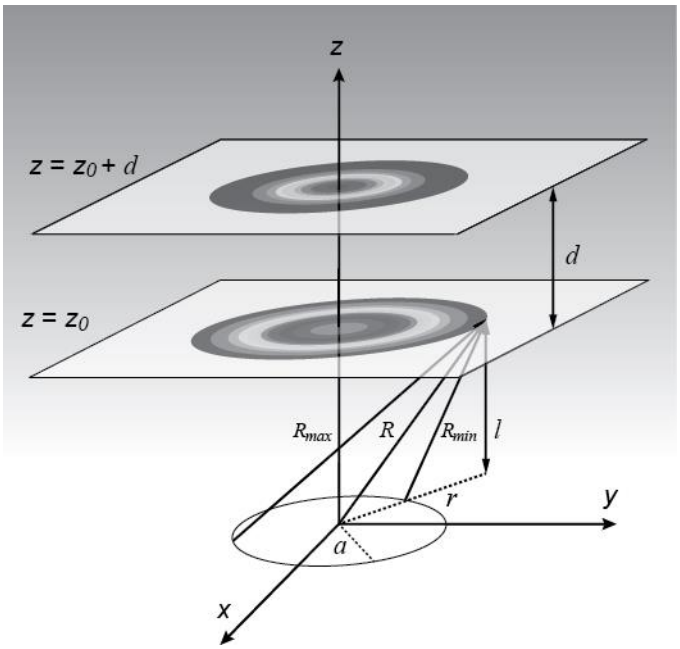


*Figure 4.7 : Geometry for the baffled piston*

The impulse response function, when $a > r$, is defined as:

$$h(R,t) = \begin{cases} 0, & ct < l \\ c, & l < ct < R_{min} \\ \frac{c}{\pi} cos^{-1}\left(\frac{(ct)^2 - l^2 + r^2 - a^2}{2r\sqrt{(ct)^2 - l^2}}\right), & R_{min} < ct < R_{max} \\ 0, & ct > R_{max} \end{cases} \quad (4.30)$$

When $a < r$, the impulse response function is defined as:

$$h(R,t) = \begin{cases} 0, & ct < R_{min} \\ \frac{c}{\pi} cos^{-1}\left(\frac{(ct)^2 - l^2 + r^2 - a^2}{2r\sqrt{(ct)^2 - l^2}}\right), & R_{min} < ct < R_{max} \\ 0, & ct > R_{max} \end{cases} \quad (4.31)$$

where $R_{min}$ and $R_{max}$ are respectively the shortest and longest distances from the observation point to the circumference of the piston:

$$\begin{aligned} R_{min} &= \sqrt{z^2 + (a-r)^2} \\ R_{max} &= \sqrt{z^2 + (a+r)^2} \end{aligned} \quad (4.32)$$

In this analysis, an under-damped sinusoidal piston motion profile is used:

$$x(t) = \mathrm{A} \cdot \sin(2\pi f t) \cdot e^{-\zeta t}, \quad (4.33)$$

$$\frac{\partial^2 x(t)}{\partial t^2} = Ae^{-\zeta t}\left(-\omega^2 sin(\omega t) - 2\zeta\omega cos(\omega t) + \zeta^2 sin(\omega t)\right), \quad (4.34)$$

where $\zeta$ is the decay coefficient. This piston displacement is considered because its derivative is known analytically, and it is representative of a plate subjected to an impulse excitation, which is the experimental case investigated in this paper. The physical parameters involved in the calculation are presented in Table 1.

*Tableau 4.2: Physical parameters*

| **Parameters** | **Symbol** | **Value** |
|---|---|---|
| Piston radius | $a$ | 0.20 m |
| Sinusoidal frequency | $f$ | 500 Hz |
| Decay coefficient | $\zeta$ | 1000 $s^{-1}$ |
| Amplitude | $A$ | 0.0001 m |
| Speed of sound | $c$ | 343 m/s |
| Air density | $\rho_0$ | 1.2 $kg/m^3$ |
| Number of samples in $x$ | $Nx$ | 231 |
| Increment in $x$ | $dx$ | 0.02 m |
| Number of samples in $y$ | $Ny$ | 231 |
| Increment in $y$ | $dy$ | 0.02 m |
| Number of time samples | $Nt$ | 482 |
| Time increment | $dt$ | 2e-5 s |
| Measuring distance | $z_0$ | 0.05 m |

### 4.7.2 **Numerical results**

Analytical and TSD-NAH results are compared for different times, positions and propagation distances. The analytical piston model is used to calculate the pressure field at $z_0 = 0.05$ m for each position of the discrete and finite planar surface described in **Error! Reference source not found.**, which is centered at $x = y = 0$. Then, this pressure field $p(z_0)$ is used as an input in the TSD-NAH algorithm. The results obtained are compared to the analytical model at $z = z_0 + d$. Four propagation distances are considered at first, that is $d = \{0.05,\ 0.20, 0.35, 0.55\}$ m. For each position of the propagation plane, the relative RMS error with respect to the time domain is calculated using:

$$\varepsilon_{RMS}(x,y) = \frac{\sqrt{\frac{\sum_t \left(p(x,y,t) - p_{th}(x,y,t)\right)^2}{N_t}}}{\sqrt{\frac{\sum_t \left(p_{th}(x,y,t)\right)^2}{N_t}}}. \quad (4.35)$$

where $p$ is the pressure field obtained with TSD-NAH and $p_{th}$ is the analytical solution.

The relative RMS is presented in Figure 4.8 for the largest propagation distance studied; $d = 0.55$ m.

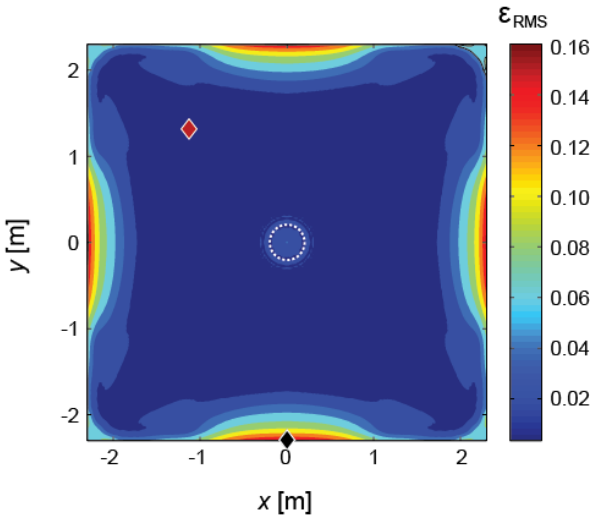


*Figure 4.8 : Relative RMS error with TSD-NAH for each position of the plane $z = z_0 + 0.55$ m. The positions of the maximum and minimum errors are marked by the pink and black diamonds, respectively. The white circle delimits the area of the piston*

The average relative RMS error in Figure 4.8 is 2.5 %, with a maximum of 16.3 % and a minimum of 0.3 %. The positions of the maximum and minimum errors are marked by the black and red diamonds respectively, corresponding to distances of around 2.37 m and 1.83 m from the center of the piston. The delimitation of the piston is represented by the white dotted circle.

The time evolution of the pressure field at the positions of the maximal and minimal spatial error and at the center of the spatial domain is presented in Figure 4.9 for the studied propagation distances in the form of dashed colored curves. The associated analytic solutions are presented in

full black lines. We observe that for most positions, the concordance with the analytical solution is high, with a relative error below 2 % at all times. The full black lines in frames (a) and (b) are not entirely visible (they appear as dotted black lines) because the analytical solutions coincide with the TSD-NAH solutions.

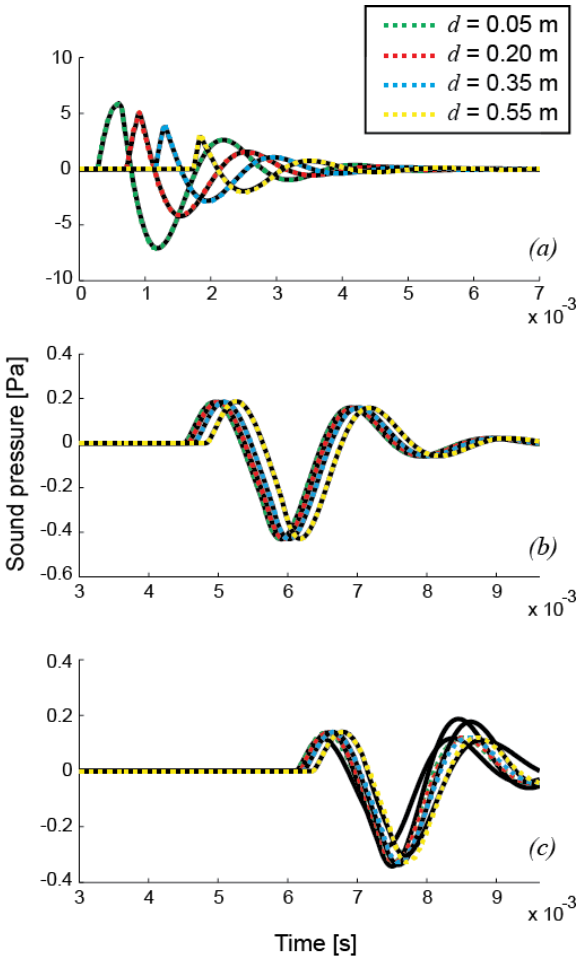


*Figure 4.9 : Comparative analysis between TSD-NAH and analytical solution to transient radiation of a baffled piston. a) Central orthogonal axis. b) Minimum error position. c) Maximum error position*

In Figure 4.8, the relative error increases at the borders of the propagation plane due to truncation errors. The impact of the truncation error can be observed in Figure 4.9c) as well; as time increases, a larger proportion of the propagating pressure field passes beyond the finite spatial boundaries of

the measurement aperture and the contribution of the truncated parts of the measured pressure field is not taken in consideration in the calculation, explaining the discrepancy with the analytic solution. This truncation error is thus more important near the boundaries. In this particular case, its contribution to the propagated pressure field is mostly positive, which leads to an underestimation of the propagated field with TSD-NAH for most cases shown in Figure 4.9c). It should be noted, however, that the causality of the signals is respected for all times and positions, with no pressure variations observed prior to the sound wave actually reaching any given position in space.

The significance of the truncation effect depends on the characteristics of the source in the time domain. In fact, since the amplitude of stationary propagative field decays with the distance as the wave travels away from the source, the amplitude of the wave is low when it reaches the boundaries compared to what it was at the source. Thus, in the convolution integral, the relative contribution of samples farther from the source is usually smaller than those of samples closer to the source, resulting in low relative error due to truncation. In the present model, however, the amplitude if the piston source decreases rapidly with time, so the amplitude of truncated samples is comparable to that at the source, causing greater truncation error.

In the examples presented, the criterion expressed in (4.24) regarding the truncation of the Green's function is respected for all propagation distances studied. Because of the good agreement of the TSD-NAH results with analytical solution, it can be concluded that (4.24) is adequate in this case. The value for coefficient $\alpha$ in (4.23) can be adapted to obtain a better compromise between the acceptable truncation errors and the computational costs, depending on the user's needs.

### 4.7.3 Comparison with other propagators

TSD-NAH results are compared to those obtained with the other propagators proposed by Hald [32] (TD-NAH), Blais [46] (TNAH) and Thomas [6] (RT-NAH). The comparative analysis of the average relative RMS error with respect to the propagation distance is presented in Figure 4.10a). The average relative RMS error is obtained by taking the mean of (4.35) with respect to $x$ and $y$. The parameters presented in Tableau 4.2 are used for all calculations, except for the spatial increment $dx$ and $dy$ and the number of samples $N_x$ and $N_y$, that are modified to 5 cm and 93 respectively. The increased spatial increment is necessary due to the higher computational requirements of some of the studied methods.

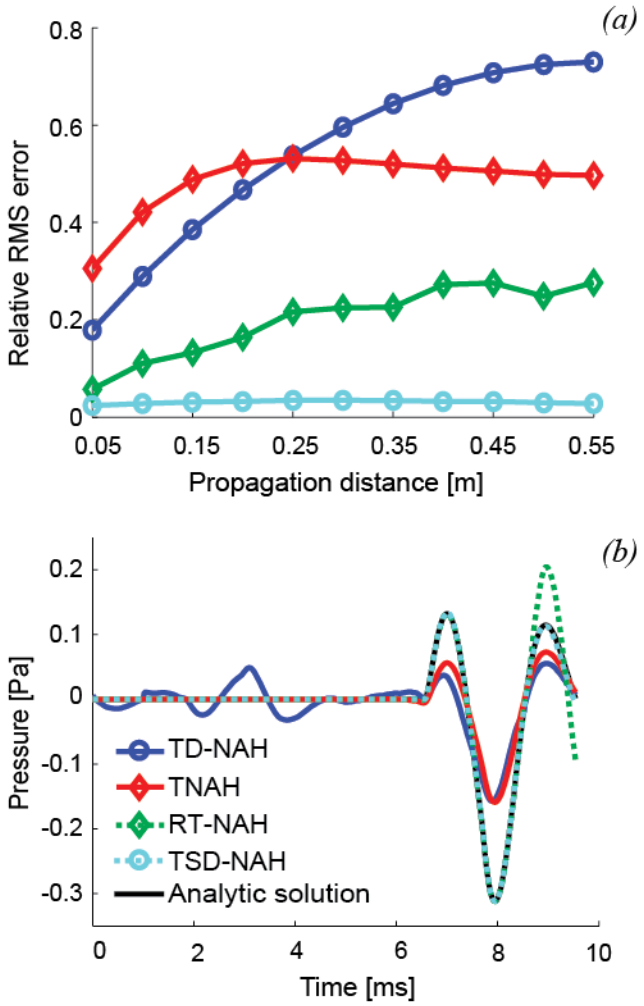


*Figure 4.10 : Comparative analysis of TSD-NAH and other TD-NAH methods. a) RMS error with respect to the propagation distance. b) Time evolution of the sound pressure at position (-1.5, 1.9) m and* $d = 0.35$ *m*

For all studied propagation distances, the relative RMS error obtained with TSD-NAH remains below 3.5 %. With the other propagators, the relative error obtained is significantly higher. At $d = 0.05$ m, the relative errors obtained with the other methods vary from around 6 % to 30 %. At $d = 0.55$ m, the relative error obtained with TSD-NAH is at least 10 times smaller than that of the other studied methods, which vary from around 28 % to 70 %. The relative error resulting from the RT-NAH method is smaller than that from TD-NAH or TNAH for all propagation distances.

An example of the time-evolution of the pressure at $x = -1.5$ m, $y = 1.9$ m and $d = 0.35$ m is presented in Figure 4.10b) for the four methods, along with the analytical solution. We observe that the TD-NAH method produces significant oscillations noticeable at the beginning of the signal, even though the signal should physically be null. The use of the numerical Laplace transform in TNAH suppresses these oscillations, however, both TD-NAH and TNAH methods underestimate the amplitude of the pressure field after the wavefront reaches the observed position. The RT-NAH method produces results that are close to the analytical solution (covering the black line that represents the analytical solution) in Figure 4.10b); however, starting from around 8.5 ms, the result diverges from the analytical solution. Also, depending on the spatial discretization, we noticed the apparition of erroneous high frequency components that modulate the signal, which were also described by Thomas [6].

It is important to note that the authors who proposed the use of these methods suggested the application of windows and filters in order to reduce leakage, which are not applied for this comparative study.

From these results, it is clear that TSD-NAH outperforms the other TD-NAH methods proposed in the literature, as it is in excellent agreement with the analytical solution. The computational requirements of TSD-NAH are higher than that of the other methods, as it requires two direct 3-D FFT and one inverse 3-D FFT over zero-padded signals. Nonetheless, with modern computers and algorithms, the calculation over a few million samples can be performed in a matter of seconds.

## 4.8 Propagation of measured pressure field from an impacted plate

In this section, propagated sound fields obtained from the application of TSD-NAH on experimental measurements are presented. The studied acoustic source is a 500 by 300 by 6 mm Plexiglas plate impacted with a metallic rod. The rod is released by and electro-magnet above the plate, and hits the plate at its center. After the impact, the plate vibrates for a few dozens of milliseconds before its movement is completely damped. The time window considered starts with the impact and lasts for around 5.3 ms. Because of the transient nature of the impact, the recorded pressure field is non-stationary.

The measurements were performed in an anechoic environment with the experimental setup is presented in Figure 4.11.

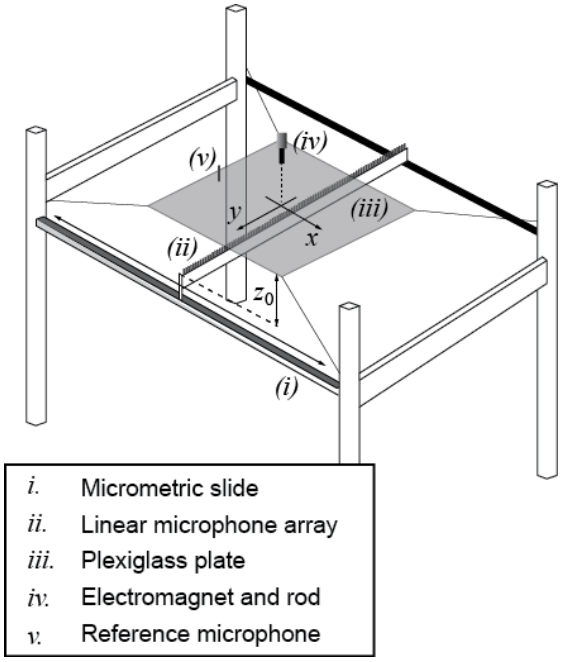


*Figure 4.11 : Experimental setup*

The measurement apparatus consists in 64 aligned class-1 microphones distanced by one cm center to center. A slide allows the translation of the line of microphones over the scanning surface (91 cm by 63 cm), with an increment of 1 *cm*. The impact is repeated for each measurement position by releasing the rod with the electro-magnet, and the measurements are phased by using a $65^{th}$ fixed reference microphone. The electro-magnet and rod setup allows a high repeatability of the plate excitation. The sampling frequency is 102.4 kHz, and the acquisition is performed with a 24-bit resolution. The plate is attached to the apparatus's frame with four metallic wires, which allows the plate to move freely in the vertical axis.

Measurements were performed in the nearfield at 65 mm from the plate. The sound field is then propagated over distances of $d = \{0.05, 0.20, 0.35, 0.55\}$ m. The results obtained in the time domain on the central orthogonal axis are presented in Figure 4.12.

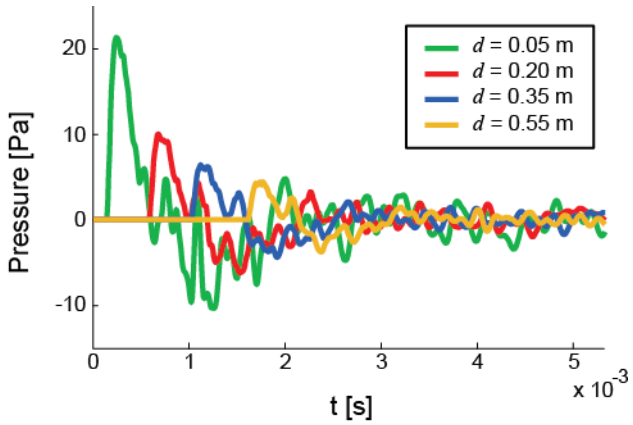


*Figure 4.12 : Forward reconstruction of the transient acoustic field. Time evolution of the pressure on the central orthogonal axis at 50, 200, 350 and 550 mm from the measurement plane*

As in the simulated case presented in Section 4.7, the forward propagated sound field have no causal errors, as the amplitude is null before the wave front reaches the propagation plane. As expected, the amplitude decreases with the propagation distance. Spatial representation of the sound field is shown in Figure 4.13 for $d = 0.55$ m and $t = 2$ ms. The pressure field is also presented along the lines $x = 0$ and $y = 0$.

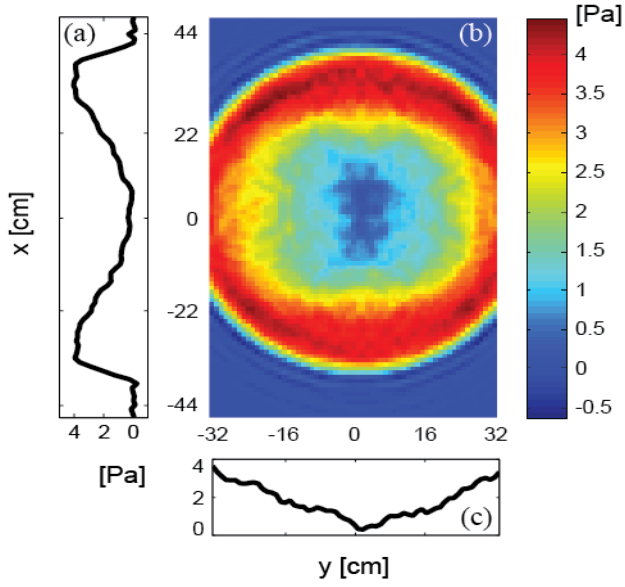


*Figure 4.13 : Spatial representation of the sound pressure field at $t = 2$ ms and $d = 55$ cm. a) Along the line $y = 0$. b) On the whole propagation plane. c) Along the line $x = 0$*

The main wave front travels radially with respect to the impact point. Other high frequency wave fronts are present due to the dispersion of the bending waves in the plate. At $d = 0.55$ m, the sound field is diffuse; reflection of the wave front at the boundaries of the plate, which is usually observed for shorter propagation distances, is unnoticeable.

## 4.9 Conclusion

Since the early 2000s, researchers have proposed formulation to express propagative non-stationary sound fields with NAH. However, leakage and wrap-around errors caused the results to diverge considerably from the analytic solutions, especially for large propagation distances.

With TSD-NAH, linear convolution of the Green's function and measured pressure field produces forward propagated sound fields without leakage or wrap-around errors. The proposed method requires sampling the analytical Green's function in the time and spatial domains. The results obtained concur well with the analytical solution for all time and space samples, and for any propagation distance. In the simulated case presented in this paper, the average error stayed below 3.5 % for propagation distances of 0.05 to 0.55 m, which is significantly lower than that of other TD-NAH methods. TSD-NAH completely suppresses causal errors such as those created by aliased replicas.

The key operation in TSD-NAH is the application of a three-dimensional linear convolution. A formula was derived to set the upper boundaries of the spatial domain of the propagator in order to avoid Green's function truncation errors. Respecting this condition implies an increased computational cost. A compromise between this cost and the generation of truncation error is proposed.

In this paper, the applicability of TSD-NAH to experimental data was shown, and the results obtained were presented for propagation distances varying from 0.05 m to 0.55 m.

In future work, application of the linear convolution formulation could be implemented to other geometries, since time domain wrap-around error is always an issue with non-stationary signals. The proposed planar formulation could also be investigated in the case of backward propagation problems.

# CHAPITRE 5 ARTICLE 2 : TIME DOMAIN NEARFIELD ACOUSTICAL HOLOGRAPHY WITH THREE-DIMENSIONAL LINEAR DECONVOLUTION

J.-M. Attendu, A. Ross, Soumis au *Journal of the Acoustical Society of America* le 28 août 2017.

## 5.1 Préface

Dans l'article 1, on a établi que la formulation proposée permet de supprimer les principales erreurs de traitement du signal, à savoir les erreurs de recouvrement et les fuites spectrales. Dans l'article 2 présenté ici, on applique cette formulation à la résolution du problème inverse, c'est-à-dire la détermination du champ acoustique entre le plan de mesure et la source. En conservant la même structure matricielle que pour le problème direct, on s'assure que la génération d'erreur de recouvrement et de fuites spectrales est évitée, ce qui permet d'atteindre les objectifs *i* et *ii* pour le problème inverse, c'est-à-dire, d'assurer la suppression du recouvrement et des fuites spectrale. Pour d'atteindre les objectifs *iii* et *iv*, qui consistent à déterminer les limites de la régularisation et des méthodes de prédiction du paramètre $\lambda$, on applique la formule de Tikhonov afin de régulariser le problème et on analyse les performances de différentes méthodes de prédiction du paramètre de régularisation optimal. On valide que la formulation proposée permet de reconstruire le champ sonore avec plus de précision que dans le cas d'une formulation se basant sur la convolution circulaire. Cependant, la formulation proposée reste imparfaite pour le problème inverse : l'utilisation du *zero-padding* engendre des erreurs de troncature propres à la formulation lorsque le champ n'est pas entièrement mesuré. On met en évidence l'impossibilité physique de mesurer le champ en entier. Néanmoins, on démontre que ces erreurs de troncature peuvent être diminuées significativement en appliquant un algorithme itératif. Malgré la complexité du problème inverse, une grande efficacité de calcul est conservée (ce qui correspond à l'objectif *v*), puisque le formalisme se base sur l'utilisation de la FFT, et que la solution à la formule de Tikhonov est connue analytiquement.

## 5.2 Abstract

In this paper, a formulation is proposed to improve the time domain reconstruction of non-stationary acoustic fields with nearfield acoustical holography. The formulation involves applying

3-D linear deconvolution using a Green's function sampled in the time and spatial domains. Because linear deconvolution does not assume periodic signals, it more appropriately describes the decaying behavior of sources that start and end at null amplitude and that radiate over a finite amount of time. The proposed method outperforms standard circular convolution-based nearfield acoustical holography by up to a factor three in relative RMS error, when compared using a transient baffled piston model, and its reconstructions remain accurate over large back-propagation distances. Furthermore, it is shown that truncation errors in linear deconvolution can be reduced by applying a 3-D patch extrapolation algorithm; however, convergence depends on the choice of an adequate Tikhonov's regularization parameter. Three methods for predicting the optimal parameter are compared: the L-curve, the generalized cross-validation and the empirical Bayesian method. It is shown that with the proposed formulation, the generalized cross-validation gives the overall best prediction for the noise levels and back-propagation distances studied.

## 5.3 Introduction

Near-field acoustical holography (NAH) is a method used to determine acoustic radiation over a 3-D space by sampling the pressure field with a 2-D microphone array located in the nearfield of the source [1, 3]. This measurement acts as a boundary condition to calculate the resulting acoustic field on other conformal surfaces by means of a three-dimensional convolution with an appropriate Green's function. Although mathematical formulations were developed for various geometries [2], this paper focuses on planar NAH: the microphone array is planar and the acoustic field is determined on planes parallel to the array, either closer to the source (the inverse problem) or further away from it (the direct problem).

Standard NAH formulations assume that the source is stationary. Consequently, the acoustic field is represented as a function of frequency $\omega$, since its frequency content is statistically constant. For non-stationary sources, the resulting field is represented in the time domain. In this paper, we assume non-stationary signals studied to have a transient nature: they start and end at null amplitude. However, we do not assume any restriction on their finite duration.

The standard NAH formulation is incompatible with non-stationary sound fields as the non-stationarity produces important signal processing errors, notably due to spectral leakage and wrap-around. For the direct problem, it was shown that these two errors can be suppressed by performing

3-D linear convolution with the Green's function sampled in the time and spatial domains (TSD) [54]. The linear convolution assumes that the convolved signals are null outside of their finite domains [22], and such assumption is representative of the non-stationary signals, assuming they have a decaying behavior in both the time and spatial domains. In this paper, we propose applying the inverse 3-D linear deconvolution to improve the solution to the inverse problem for non-stationary acoustic sources.

Time domain unidimensional linear deconvolution was previously applied to non-stationary NAH with real-time nearfield acoustical holography (RT-NAH) [34, 36], which is based on the use of the time-wavenumber Green's function [35]. The plane-wave superposition method proposed by Geng, Zhang and Bi also uses the time-wavenumber Green's function for reconstructing non-stationary sources [55, 38]. Blais and Ross proposed the use of the numerical Laplace transform in the time domain to reduce wrap-around effects [8]. Wu also used the numerical Laplace transform to reconstruct sound fields using the Helmholtz least-squares method [39]. However, the inverse Laplace operator is subject to the amplification of numerical errors due to the application of exponential window [47].

The particularity of the method presented in this paper is that the linear deconvolution is applied to all three dimensions, including the spatial domain. Errors due to the use of inverse circular convolution in the spatial domain are thus avoided. Additionally, the operations involved are based on the use of the fast Fourier transform (FFT) which makes the method computationally efficient.

Extension of the iterative patch extrapolation process [16, 17, 18, 19] to the time domain for non-stationary sources and its applicability for reducing truncation errors is investigated. We show that it can be used successfully in conjunction with the proposed Green's function.

Finally, since the backward reconstruction is an ill-posed problem, Tikhonov regularization is applied to reduce the negative effect of evanescent components amplification [9, 56]. The effectiveness of different methods for predicting the optimal regularization parameter in Tikhonov's formula for the proposed time-space Green's function is assessed. The methods include the empirical Bayesian method which has shown promising performance in recent work [30, 37], along with the L-curve and Generalized cross-validation methods [9, 56]. These methods do not require prior knowledge of the noise variance.

The present paper thus brings the following three main contributions:

1. A formulation for inverse non-stationary NAH with linear deconvolution (inverse 3-D Toeplitz operator) and the TSD Green's function which reduces the reconstruction error is presented;
2. The applicability of the iterative patch extrapolation algorithm to the TSD method in order to reduce the effect of the truncated measurements is discussed;
3. The comparative analysis of three methods to predict the optimal Tikhonov regularization coefficient with the proposed formulation is presented.

Each of these three objectives constitutes one of the sections of this paper.

## 5.4 Non-stationary nearfield acoustical holography

In this section, we present the theory behind TSD non-stationary near-field acoustical holography along with the different signal processing problems caused by the deconvolution of finite and discrete non-stationary signals. We also present a comparative analysis of the proposed and standard NAH methods applied to time domain reconstruction.

The planar geometry considered is illustrated in Figure 5.1.

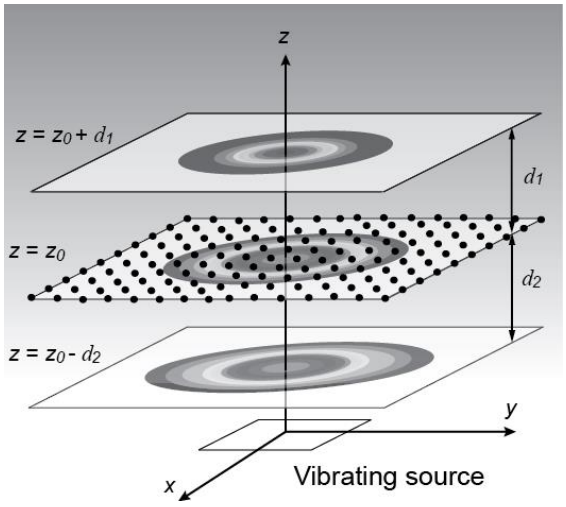


*Figure 5.1 : Geometry for planar near-field acoustical holography*

A finite baffled source is represented by the rectangular surface centered at $z = 0$ on the *x-y* plane. The microphone array, consisting of a rectangular grid of uniformly spaced microphones (represented by dots) is located on the parallel plane $z = z_0$. The focus of this paper is backward

reconstruction: the acoustic field is determined on a parallel plane located between the measurement plane and the source. This is illustrated in Figure 5.1 at $z = z_0 - d_2$; a forward reconstruction plane is shown at $z = z_0 + d_1$, where $d$ is the propagation distance in either direction ($d > 0$).

The mathematical relation between sound pressures on two parallel planes is given by the three-dimensional convolution with respect to the time ($t$) and the spatial ($x$ and $y$) domains, as expressed in Eq. (5.1):

$$p(t, x, y, z + d) = \iiint_{-\infty}^{\infty} p(t', x', y', z) \cdot g(t - t', x - x', y - y', d) dt' dx' dy' \tag{5.1}$$

The pressure field expressed inside the convolution integral $p(z)$ is the one closer to the source, and $g$ is the time and spatial domains (TSD) Green's function expressed in Eq. (5.2) [12, 7]:

$$g(t, x, y, d) = \frac{d}{2\pi c}\left(\frac{c}{R^3} + \frac{1}{R^2}\frac{\partial}{\partial t}\right)\delta(t - R/c). \tag{5.2}$$

The phase velocity is represented by $c$ and $R = \sqrt{x^2 + y^2 + d^2}$ is the sound propagation distance. By using Euler's equation [13] on either the measured pressure or on the Green's function, particle velocity, intensity or acceleration can also be determined.

### 5.4.1 **Discrete convolution**

In order to obtain an inverse formulation for Eq. (4.1) and to deal with the finite and discrete nature of the measured signal, the convolution integral in Eq. (4.1) is expressed in a matrix equation form:

$$p_m = G \cdot p_s. \tag{5.3}$$

In this case, vector $p_m$ expresses the measured pressure field at $z = z_0$ and vector $p_s$ represents the unknown pressure field near the source at $z = z_0 - d$. The goal of the deconvolution problem is to determine $p_s$, which is obtained by inverting $G$, the transformation matrix which encompasses the 3-D convolution operation with the Green's function. Matrix $G$ can take different forms depending on the assumptions regarding the discrete convolution.

There are two standard ways for expressing discrete convolution:

i. Either we assume periodicity of the finite convolved signals,

ii. or we consider the finite signals to be null outside of their domains of definition.

These two assumptions lead to two different convolution results, namely the circular (or cyclic) and linear convolutions [22]. Wrap-around error occurs when circular convolution is applied even though the acoustic field is not periodic.

These two types of convolutions can be interpreted in the context of NAH. In the time domain, if the radiated sound field is stationary, circular convolution should be applied because the periodic repetitions are representative of the signal outside of the measured domain. For non-stationary acoustic sources, assuming that the amplitude of the source decreases to zero with time (which is the case of most sources when the measurement time is long enough), the linear convolution assumption is appropriate. In the spatial domains $x$ and $y$, it is assumed that the source is of finite dimension and is smaller than the microphone array. Since the radiated pressure field decreases with propagation distance, the sound field's amplitude is concentrated mostly near the source; consequently, spatial linear convolution is adequate for both stationary and non-stationary signals.

Therefore, for non-stationary sources with decreasing amplitude, three-dimensional linear deconvolution should be applied.

### 5.4.2 Computation of the linear deconvolution

The three-dimensional linear convolution matrix $G$ takes the form of a Toeplitz-block Toeplitz matrix. For a typical NAH setup, this matrix is composed of a very large number of elements (more than the square of the product of the number of elements in each of the three dimensions) which makes its inversion or singular value decomposition impractical. There are means to iteratively calculate the inverse of Toeplitz-block Toeplitz matrix with improved efficiency [57, 58], but their consideration is beyond the scope of this paper. Instead, we embed the Toeplitz matrix in a bigger block-circulant matrix which the inverse can be determined very efficiently by using the 3-D FFT [10]. Such formulation is equivalent only if the whole pressure field is fully measured; otherwise, truncation errors are introduced, as discussed in Section 5.5.

In practice, deconvolution is implemented with the FFT as illustrated in Figure 5.2.

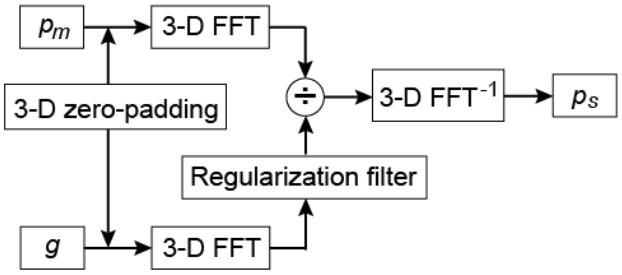


*Figure 5.2 : Operations involved in the 3-D deconvolution with the circularization method*

Three-dimensional zero-padding is applied to the TSD Green's function: for each dimension, zero values are added at the end of the signal to double its length, thus multiplying the total number of elements by eight. Zero padding is also applied to the measured pressure field to fit the size of the Green's function if necessary. Three-dimensional FFT is then applied to both signals. A regularization filter derived from Tikhonov's formula is applied to the Green's function in the Fourier domain in order to stabilize the amplification effect of noisy evanescent components [9]. The pressure field is divided by the filtered Green's function in the Fourier domain, and an inverse 3-D FFT is applied to obtain the reconstruction in the time and spatial domains.

### 5.4.3 Tikhonov regularization

Since deconvolution is an ill-posed problem, regularization must be applied in order to limit noise amplification. The standard method for regularization in NAH is to find the regularized solution $\widehat{p_s}$ that minimizes Tikhonov's cost function:

$$argmin\left(\left\|Gp_s - p_m\right\|^2 + \lambda \cdot \|p_s\|^2\right), \tag{5.4}$$

where $G$ is the convolution matrix, $\lambda$ is the regularization parameter (a Lagrange multiplier) and $\|.\|$ represents the $\ell^2$ norm. The solution to Eq. (5.4) is:

$$\widehat{p_s} = (G^T G + \lambda)^{-1}(G^T p_m). \tag{5.5}$$

where $G^T$ is the transpose of $G$ and $\widehat{p_s}$ is the regularized solution. Considering the circulant form of the zero-padded linear convolution matrix, and considering that such matrix is diagonalized by

the discrete Fourier transform (DFT) operator [19], the regularized solution $\widehat{p_s}$, in this case, is equivalently expressed as:

$$\widehat{p_s} = \mathcal{F}^{-1}\left\{\frac{|\mathcal{F}\{g\}|^2}{|\mathcal{F}\{g\}|^2 + \lambda} \cdot \frac{\mathcal{F}\{p_m\}}{\mathcal{F}\{g\}}\right\}, \tag{5.6}$$

with $\mathcal{F}$ and $\mathcal{F}^{-1}$ symbolizing the direct and inverse 3-D DFTs. Eq. (5.6) has the advantage that the Green's function $g$ and its FFT can be calculated efficiently with a desktop computer.

The left-hand side fractional term in Eq. (5.6) is a low pass Wiener filter, and the regularization parameter $\lambda$ is analogous to its cut-off frequency. Dividing the measurement by the band-limited Green's function amplifies the high frequency components; this amplification becomes more important as the back-propagation distance increases. The Wiener filter thus limits the amplification; however, it also affects the amplitude and the shape of the solution.

Tikhonov regularization is a standard method in NAH: it is a convenient method because its solution is expressed in a closed form and it can be computed efficiently by using the FFT. However, this solution to the regularization problem is not optimal.

### 5.4.4 **Leakage errors**

With time domain NAH, if the analytical Green's function is sampled in the Fourier domain as it is the case with stationary NAH formulations, spectral leakage is generated when the reconstruction $p_s$ is transformed back into the TSD domain. These numerical errors take the form of replicas and high frequency ripples due to the application of the inverse Fourier transform on the discrete finite Green's function. Sampling the analytical Green's function directly in the TSD prevents generating such errors.

To illustrate this, we compare the TSD Green's function with the corresponding Green's function obtained from the inverse 3-D DFT of the analytical Fourier domain function [3], as presented in Figure 5.3(a) and (b). This can be thought of as the solution to the reconstruction of an impulsive point source. In order to highlight the presence of leakage errors, the logarithmic scale of the absolute value is presented in Figure 5.3 (b), whereas Figure 5.3 (a) is presented in linear scale since the amplitude of the TSD Green's function is mostly null.

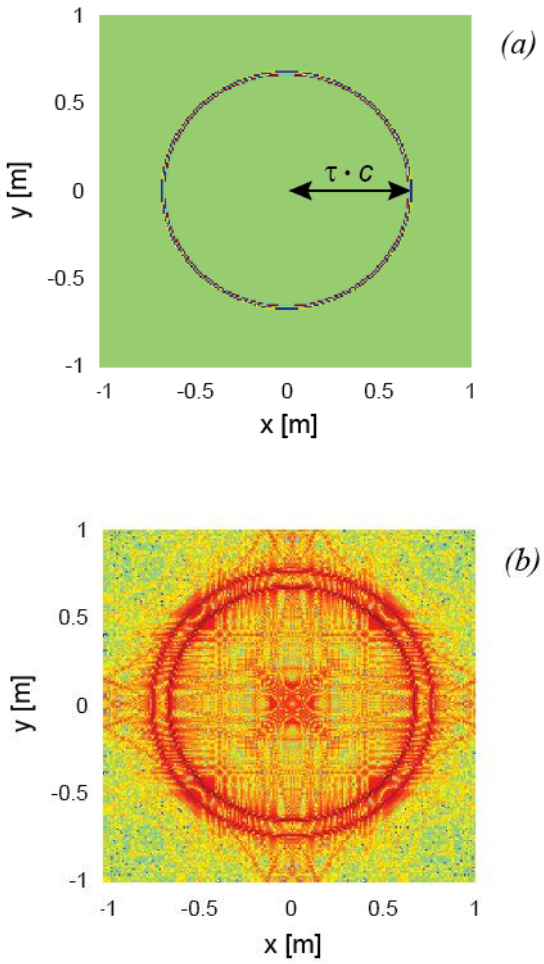


*Figure 5.3 : Amplitude of the Green's functions expressed in the spatial domain at $\tau = 2$ ms. a) TSD-NAH formulation. b) Obtained from the inverse 3-D DFT of the analytical expression in the Fourier domain (logarithmic scale). Adapted from Fig. 2 in [54]*

In Figure 5.3 (a), the TSD Green's function is a series of Dirac deltas positioned at a constant distance from the center. The radius corresponds to time $\tau$ multiplied by sound velocity $c$. The amplitude of $g$ is null at all other positions, as expected.

In Figure 5.3 (b), most of the amplitude is also distributed at a constant radius with respect to the center, which corresponds to the spatial shift presented in Figure 5.3a). The second concentric distribution with smaller radius corresponds to aliased replicas caused by leakage in the time domain. Significant ripples are distributed over the whole domain, particularly near the center. Spatial replicas due to the Fourier domain discretization are also noticeable.

Along with suppressing wrap-around errors, the use of the TSD Green's function is preferable for time domain reconstruction, as it avoids generating such ripples and replications.

### 5.4.5 Comparative analysis of the reconstruction error

In order to quantify the benefits of using the TSD method for solving the inverse problem, we present a comparative study of the reconstructions obtained with the proposed method and the Fourier domain Green's function expressed as [2]:

$$g(\omega, k_x, k_y, d) = e^{id\sqrt{\frac{\omega^2}{c^2}-k_x^2-k_y^2}}, \tag{5.7}$$

where $k_x$ and $k_y$ are the magnitudes of the wave-number vectors projected on the $x$ and $y$ axis, and $\omega$ is the angular frequency. Both Green's functions (Eqs. (2) and (7)) are equivalent in their continuous form, as one is the analytical 3-D Fourier transform of the other. However, when applied to the discrete finite formulation, signal processing errors generated with both methods differ. For the Fourier based operator, in addition to spectral leakage, wrap-around error is generated due to the assumption that Eq. (5.7) is periodic. Adding zero-padding solely to the measurement $p_m$ as proposed in the literature does not suppress wrap-around, but only reduces its effect.

A non-stationary baffled piston model [45] is used to generate the pressure fields on planes $z = z_0 - d$ and $z = z_0$. An overview of the piston model is presented in Appendix A, along with the physical parameters considered. The reconstruction is calculated on plane $z = z_0 - d$ from the values on plane $z = z_0$ with both the standard Fourier-based NAH and the TSD methods. The relative RMS error is calculated for each method by comparing the reconstruction $p_s$ with the result predicted by the piston model $p_{th}$ using Eq. (5.8):

$$\varepsilon_{RMS} = \frac{\sqrt{\sum_{xyt}(\widehat{p_s} - p_{th})^2}}{\sqrt{\sum_{xyt} p_{th}^2}}. \tag{5.8}$$

Eq. (5.6) is used to calculate the reconstruction $p_s$ obtained with regularization parameter $\lambda$ varying from $10^{-5}$ to $10^{2}$. For both methods, the optimum regularized reconstruction $\widehat{p_s}$ considered is the

solution $p_s$ that minimizes $\varepsilon_{RMS}$. Although the optimal regularization parameter is considered here, means for accurately predicting such value are discussed in Section 5.6.

The spatial discretization is $dx = dy = 1$ cm for a 1 meter wide square aperture, and the time domain increment is $dt = 0.01$ ms over a period of 5 ms. The relative RMS errors obtained from Eq. (8) are presented in Figure 5.4 for back-propagation distances of 3 cm to 20 cm (i.e. the "measurement plane" is 3.5 to 20.5 cm from the source, and the reconstruction plane is 5mm from the source).

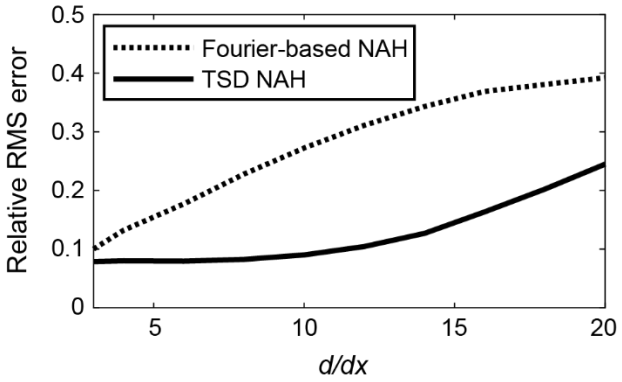


*Figure 5.4 : Relative RMS error obtained with the Fourier-based NAH and the TSD methods with respect to the propagation distance*

The back-propagation distances presented are typical of NAH, since the microphone array is located in the near field of the source.

We observe that for all back-propagation distances studied, the relative RMS error for the TSD NAH formulation is significantly smaller than that of the Fourier-based NAH formulation. For $d = 12$ cm, the difference in relative RMS error is maximal at 20.7 % and the relative error for TSD NAH is a third of that of the circular deconvolution method. For $d = 20$ cm, the difference in relative RMS error is about 14.5 %. For $d = 3$ cm, both methods lead to low relative RMS errors, since the measurement is similar to the theoretical reconstruction and the deconvolution does not significantly alter the measurement.

For shorter back-propagation distances (up to about 8 cm), the relative error obtained with the TSD method remains at about 8 %, and the inverse problem has a higher numerical stability as the

optimal value of $\lambda$ is low. The components filtered by the Wiener filter mostly contain numerical noise and the frequency content of the acoustical signal remains unaffected.

Beyond $d = 8$ cm, the relative RMS error increases with respect to the back-propagation distance. This is partly due to the fact that the optimal Tikhonov's solution requires more severe regularization: for $d = 20$ cm, the optimal regularization parameter is 5 times larger than for $d = 4$ cm (although optimal value for the regularization parameter does not increase linearly with respect to $d$). The regularization filter prevents amplification of noise components, but it also affects signal components, thus contributing to the discrepancy with the solution.

Furthermore, truncation errors also increase with back-propagation distance. By placing the microphone array farther away from the source, a greater proportion of the sound field with respect to space and time is not captured by the finite measurement aperture. The truncated components induce important errors in the reconstruction, as discussed in the next section.

In Figure 5.5, two examples of the time domain reconstruction obtained with both methods are presented along with the theoretical solution for $d = 5$ cm and $d = 18$ cm. For comparison purposes, the theoretical under-damped sinusoidal motion of the piston is also shown.

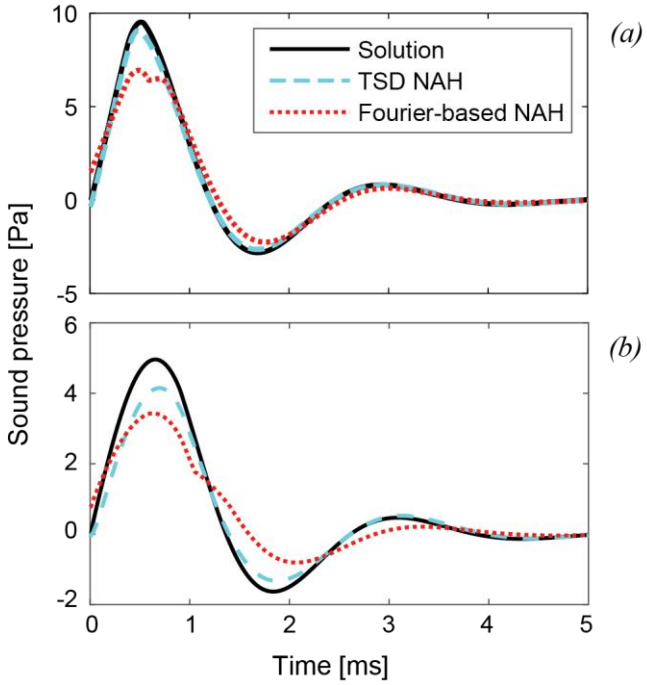


*Figure 5.5 : Evolution of the pressure field in the time domain. (a)* $(x, y) = (0,0)$ *m and* $d = 0.05$ *m. (b)* $(x, y) = (0.08,0.08)$ *m and* $d = 0.18$ *m*

In Figure 5.5 (a), the time domain reconstruction is presented at the centre of the aperture. Here, TSD NAH only slightly underestimates the amplitude of the solution. The discrepancy is more important for the larger back-propagation distance in Figure 5.5 (b), presented at x = y = 8 cm, since effects of the regularization filter and truncation errors are more important as the back-propagation distance increases. In addition, the discrepancy increases with $x$ and $y$ as truncation effects are more important near the boundaries of the aperture.

With the Fourier-based method, the discrepancy with the solution is more considerable, especially for the first rise in amplitude: there is an inflection in the curve, which corresponds to a duplication of the lobe. This wrap-around error is caused by the contribution of the periodical replicas in circular deconvolution. Such effect is also observable in Figure 5.5 (b), but in that case, the second lobe is shifted to the right. We observe that the magnitude of the shift increases with the back-propagation distance. This could be explained by the increase in travelling time between the reconstruction domain and the replications with respect to $d$. Additionally, leakage effect in the Green's function contributes to increasing the discrepancy, although high frequency leakage is

attenuated by the application of the Wiener filter. Finally, in both graphs, the Fourier-based reconstruction does not start at zero. This does not respect the physical causality of the problem, as the signal should be null before $t = 0$.

Errors obtained with Fourier-based NAH are generally observed at other back-propagation distances and spatial positions. Such discrepancies are explained by the fact that the mathematical model for the circular convolution with the Fourier based operator is inadequate: it does not correspond to the physical assumptions of the problem as it assume periodicity. With the TSD method, we can control the error by reducing the discretization and increasing the size of the aperture; however, with the Fourier-based operator, errors inherent to the convolution model are generated and cannot be avoided.

Finally, the effect of wrap-around error and leakage occur in both the time domain and in the spatial domains. Consequently, the full 3-D linear deconvolution must be used. In that sense, use of the time and wave-number Green's function as proposed by many in the context of non-stationary NAH [35, 34] does suppress such errors in the time domain, however, since the circular deconvolution is still used in the $x$ and $y$ domains, effects due to the spatial replicas still occur. To the author's knowledge, the time-wave number operator has only been applied to back-propagation distances in the order of a few times the spatial discretization; no examples of accurate reconstruction for larger back-propagation distance, and consequently more prone to wrap-around error and ill-posedness, were presented in the literature.

## 5.5 Truncation error and patch algorithm

With the proposed method, the linear deconvolution is obtained by embedding the Toeplitz-block Toeplitz matrix $G$ in a larger block-circulant matrix. This implies that the measured signal encompasses the full acoustic field. However, since sampling is done over a finite domain, a truncation error is introduced. In this section, we demonstrate that the signals cannot be fully captured both in the time and spatial domains, and therefore that truncation is inevitable. We formalize truncation error in the context of TSD NAH, and we show that the patch algorithm can be applied in 3-D as a means to reduce these errors.

### 5.5.1 **Truncation errors**

The Toeplitz-block Toeplitz matrix is embedded in a block-circulant matrix that can be diagonalized using the 3-D FFT. In return for its high computational efficiency, the block-circulant based linear deconvolution involves a computational degradation of the solution if the measured pressure field is truncated. This is illustrated with a simplified 1-D example: we suppose that source $p_s$ and propagator $g$ are known and that both have $N$ values. The full propagated signal $p_m$ is obtained by calculating the linear convolution of $p_s$ and $g$. Zero-padding is applied to both $p_s$ and $g$ before embedding the propagator in a $2N$ by $2N$ circulant matrix:

$$C = \begin{bmatrix} G_1 & G_2 \\ G_2 & G_1 \end{bmatrix}. \tag{5.9}$$

$G_1$ and $G_2$ are the first and second halves of the $2N$ by $N$ Toeplitz linear convolution matrix which takes the following form:

$$[G_1 \quad G_2]^T = \begin{bmatrix} g_1 & g_2 & g_3 & \cdots & g_N & 0 & 0 & 0 & \cdots & 0 \\ 0 & g_1 & g_2 & g_3 & \cdots & g_N & 0 & 0 & \cdots & 0 \\ 0 & 0 & g_1 & g_2 & g_3 & \cdots & g_N & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & 0 \\ 0 & \cdots & 0 & g_1 & g_2 & \cdots & g_{N-1} & g_N & 0 & \vdots \\ 0 & \cdots & 0 & 0 & g_1 & g_2 & \cdots & g_{N-1} & g_N & 0 \end{bmatrix}^T, \tag{5.10}$$

where superscript $T$ symbolizes the transpose operation. The inverse of the circulant $C$ is obtained from the FFT of the zero-padded Green's function, since all circulants are diagonalized by the discrete Fourier transform (DFT) matrix [19].

The full linear convolution result is given by:

$$\begin{bmatrix} G_1 & G_2 \\ G_2 & G_1 \end{bmatrix} \cdot \begin{bmatrix} p_s \\ 0 \end{bmatrix} = \begin{bmatrix} p_{m1} \\ p_{m2} \end{bmatrix}, \tag{5.11}$$

or equivalently,

$$\begin{aligned} p_{m1} &= G_1 \cdot p_s, \\ p_{m2} &= G_2 \cdot p_s, \end{aligned} \tag{5.12}$$

with $p_{m1}$ and $p_{m2}$ being the first and last $N$ values of the linear convolution respectively, and the $0$ symbol represents the sequence of $N$ null values added to $p_s$.

We now consider the truncated inverse problem in which the last $N$ values of $p_m$ are not measured. We replace the second half of $p_m$ with null values and the inverse system becomes:

$$\begin{bmatrix} G_1 & \mathrm{G}_2 \\ \mathrm{G}_2 & G_1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} p_{m_1} \\ 0 \end{bmatrix} = \begin{bmatrix} p_s + \epsilon_1 \\ \epsilon_2 \end{bmatrix}, \tag{5.13}$$

where $\epsilon_1$ and $\epsilon_2$ are truncation errors. Because of its block-matrix structure, the inverse of $C$ can be obtained with respect to its block components [59]:

$$C^{-1} = \begin{bmatrix} G_1^{-1} + G_1^{-1}\mathrm{G}_2 S^{-1}\mathrm{G}_2 G_1^{-1} & -G_1^{-1}\mathrm{G}_2 S^{-1} \\ -S^{-1}\mathrm{G}_2 G_1^{-1} & S^{-1} \end{bmatrix}. \tag{5.14}$$

The Schur complement $S$ is given by:

$$S = G_1 - \mathrm{G}_2 G_1^{-1}\mathrm{G}_2. \tag{5.15}$$

Therefore, the truncation errors are:

$$\epsilon_1 = G_1^{-1}\mathrm{G}_2 S^{-1}\mathrm{G}_2 G_1^{-1} \cdot p_{m_1} = G_1^{-1}\mathrm{G}_2 S^{-1} \cdot p_{m_2}, \tag{5.16}$$

$$\epsilon_2 = -G_1^{-1}\mathrm{G}_2 S^{-1} \cdot p_{m_1} = -S^{-1} \cdot p_{m_2}. \tag{5.17}$$

If the truncated portion of the signal is not significant ($p_{m_2} \to 0$), according to (5.16) and (5.17), both $\epsilon_1$ and $\epsilon_2$ tend to zero. Note that in this example, if the inverse of matrix $G_1$ were computable, the reconstruction $p_s$ could be obtained directly from the product of $G_1^{-1}$ and $p_{m_1}$: it is the embedding in the circulant matrix that causes the truncation error.

In practice, since the linear deconvolution is applied to all three dimensions, truncation errors are generated in each truncated domain $x$, $y$ and $t$.

### 5.5.2 Duality of the truncation in the time and spatial domains

Because sampling is performed in both the time and spatial domains, truncation of the full field is inevitable. In fact, there is duality between the two domains: for a finite aperture, if the signal is

fully measured in the time domain at each position (i.e. the signal is measured until it decreases to zero), the spatial domain will be truncated, and vice versa.

To demonstrate this, we suppose a point source located at $x = y = 0$ that radiates from $t = 0$ to $t = \tau$ seconds. We measure the signal with a finite square aperture of diagonal $2 \cdot r_{max}$ over a period of $t_{max}$.

In order for the signal to be fully recorded in the time domain, $t_{max}$ must be greater than the duration $\tau$, in addition to the time taken by the wave front to reach the position of the microphones farthest from the source:

$$t_{max} \geq \tau + \frac{r_{max}}{c}. \tag{5.18}$$

To avoid truncation in the spatial domain, the shortest distance between the centre and the perimeter of the square aperture (corresponding to its half-width $r_{max}/\sqrt{2}$) must be larger than the largest travelling distance of the wave front $t_{max} \cdot c$:

$$\frac{r_{max}}{\sqrt{2}} \geq t_{max} \cdot c. \tag{5.19}$$

It is not possible to have $t_{max}$ and $r_{max}$ that respect both equations (5.18) and (5.19). This duality is illustrated in Figure 5.6.

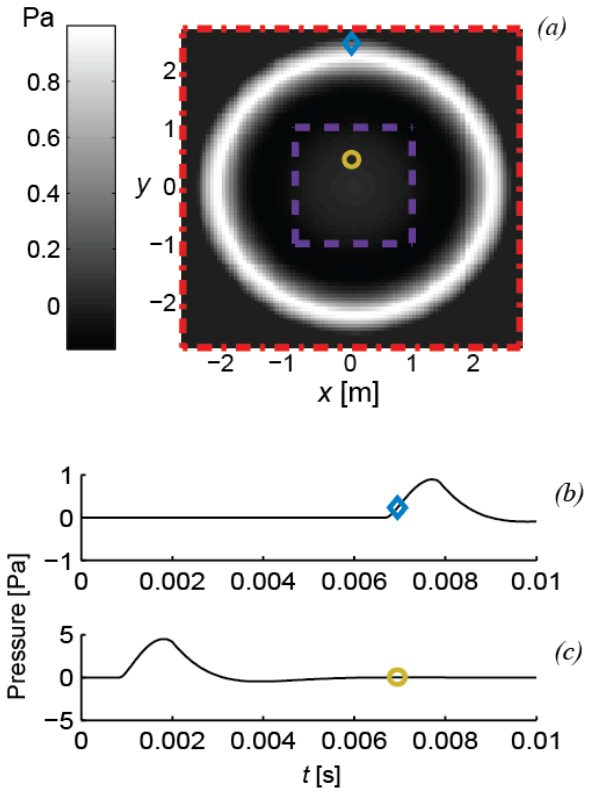


*Figure 5.6 : Space-time duality. (a) Spatial-domain truncation. (b) Time domain truncation at (0,2.5) m. (c) Time domain truncation at (0,0.5) m*

In Figure 5.6 (a), a spatial snapshot of the sound field at the source plane is presented at the time sample $t_{max} = 7$ ms. The red and purple dashed squares outline the boundaries of two finite spatial domains. The red boundaries respect the spatial condition expressed in Eq. (5.19). However, in the time domain, the signal is truncated for various positions. For example, for the position indicated by the blue diamond, stopping the measurement at $t = 7$ ms causes truncation of the signal, as represented in Figure 5.6 (b).

On the other hand, for positions closer to the source, such the one indicated by a yellow circle, the time domain condition is respected, as illustrated in Figure 5.6(c). However, a spatial aperture that encompasses only those positions, such as that represented by the purple boundaries in Figure 5.6(a), implies that spatial truncation occurred when the wave front travelled through the boundaries.

Truncation of the measurement is therefore inevitable. However, since the amplitude of the pressure decreases radially, the effect of the truncation can be reduced by selecting a 3-D window

that respects the time domain condition, and that is wide enough in order for the spatial truncation to occur at low amplitudes.

### 5.5.3 Truncation of the Green's function

The time-space duality also affects the Green's function. However, its amplitude is subject to similar radial decrement in the TSD, and the spatial window can be selected accordingly. Since the Green's function is known analytically, its span is not limited by the measurement aperture and can be as wide as necessary, with the drawback of increased computational costs. In previous work, it was shown that for a fixed aperture size, the truncation of the Green's function approximately increases linearly with propagation distance $d$ [54].

### 5.5.4 Patch extrapolation

Since having a large measurement aperture is not always possible, another way of dealing with truncation is to consider patch extrapolation [60, 61, 62, 63]. The idea behind this method is to iteratively calculate the propagated field at the measurement plane using the reconstruction as an input for the direct problem. For each iteration, the data on the propagation field is replaced with the actual measurement, and the data extrapolated outside of the measurement window is kept. The extrapolated data compensates for the truncation, and the reconstructed field becomes increasingly accurate with each iteration.

The mathematical relation between the consecutive $(i-1)^{th}$ and $i^{th}$ iteration is the following:

$$\begin{bmatrix} p_s^i \\ \zeta_i \end{bmatrix} = C^{-1} \cdot \begin{bmatrix} p_{m_1} \\ G_2 \cdot p_s^{i-1} \end{bmatrix}, \tag{5.20}$$

where $p_s^i$ is $i^{th}$ iteration of the reconstructed field and $\zeta_i$ are the values of the reconstructed pressure field outside of the source's location. Since the source is baffled, the values of $z_i$ should tend to zero as $p_s^i \to p_s$ (similarly as $\epsilon_2 \to 0$). When the reconstruction is performed for $d \neq z_0$, it is required that the field's amplitude is close to zero outside of the finite domain, a condition only obtained with wider spatial boundaries as discussed above. It is also required that measurement $p_{m_1}$ be at least as large as the baffled source $p_s$.

This algorithm is important in NAH because its iterative process has been applied to many sound reconstruction formulations, such as SONAH, HELS and ESM [41, 42, 64]. To the authors' knowledge, convergence of the patch algorithm has not been demonstrated rigorously, although it was discussed by Hansen [10]. Saijyou et al. also presented an insightful demonstration that supports convergence in the context of NAH [63]. Convergence can be obtained when applied to the proposed linear convolution formulation, although the algorithm is sensitive to the regularization parameter used. We present an example of this convergence by applying the patch algorithm to the reconstruction of the sound field radiated by the circular baffled piston model detailed in Appendix A, using the proposed 3-D linear deconvolution formulation.

The pressure field is determined on three square apertures with lengths varying from 0.2 m to 1.2 m. These fields are back-propagated using the TSD method and the reconstructions are compared to the solution given by the piston model. Eq. (5.4) is used to calculate the relative RMS error. These results are represented by the curves in Figure 5.7 for $d = \{4, 20, 36\}$ cm, with $dx = 4$ cm, $dt = 0.02$ ms and a period of 5 ms.

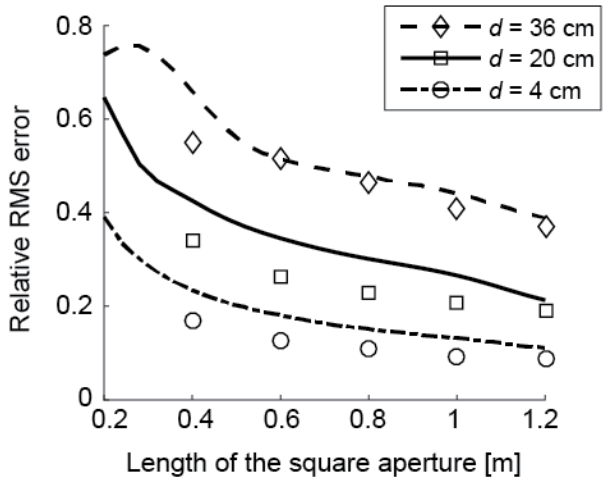


*Figure 5.7 : Relative RMS error obtained with regular TSD NAH (lines) and patch algorithm (markers) with respect to the length of the aperture and the back-propagation distance*

As expected, we observe that the relative RMS error decreases as the size of the aperture increases due to the reduced truncation errors. The global behavior of the truncation error appears to be approximately the same with respect to aperture size for all back-propagation distance. The error increases with back propagation distance since the ill-posedness of the problem increases with $d$,

and also because truncation of the field is more important when the measurement aperture is farther from the source.

The relative RMS errors obtained with patch algorithm are represented by circle ($d = 4$ cm), square ($d = 20$ cm) and diamond ($d = 36$ cm) markers. In each case, the original aperture considered is half the dimensions of the final extrapolated domain, i.e. a quarter of its area.

The effectiveness of the patch algorithm is beyond expectations: in all cases studied, the resulting relative RMS error obtained with the extrapolated aperture is below or equal to that of the reconstruction obtained from the full aperture. In other words, the reconstruction produced with the iteratively extrapolated measurement is more accurate than that of the full measurement without extrapolation. Although we have not verified this, a realistic explanation for the performance of the patch algorithm is that regularization is applied cumulatively at each iteration, and this process filters more adequately the noised components of the measured field. To support this, it was shown by Williams [56] that embedding Wiener filters (two iterations in his case) results in a steeper filter that reduces the negative effect of regularization on both propagative and evanescent waves components.

The performance of the patch extrapolation algorithm depends on the regularization parameter used. For low values of $\lambda$, regularization is inadequate and the contribution of amplified noisy components adds up at each iteration: consequently, the patch algorithm diverges. For high values of $\lambda$, the reconstruction quickly iterates to zero, as important filtering of the signal occurs at each iteration. Methods for predicting the optimal regularization parameter are presented in the next section.

## 5.6 Predicting the optimal Tikhonov's regularization parameter

Regularization is a fundamental operation in the inverse NAH problem, as it prevents amplification of noisy evanescent components [56]. In this section, we apply the standard Tikhonov regularization to the linear convolution-based NAH. This requires determination of the regularization parameter $\lambda$ in Eq. (5.6). Three different methods for determining $\lambda$ are compared, that is: the L-curve [25], the generalized cross validation [29] and the empirical Bayesian approach [30]. A common ground to the studied methods is that they do not require prior knowledge of the

noise variance. The parameter values obtained are analysed with respect to the propagation distance and the noise level.

### 5.6.1 **L-curve**

The regularization parameter is obtained with the L-curve method by finding a compromise between the residual norm $\left\| g * \widehat{p_s}(\lambda) - p_m \right\|$ and the norm of the solution $\|\widehat{p_s}(\lambda)\|$. Plotting the logarithm of these two quantities as a function of $\lambda$ produces an "L" shaped curve, and an optimal regularization parameter is obtained at the maximum curvature.

The curvature is obtained by minimizing the cost function [25]:

$$\mathrm{J}_{Lcurve} = -2\frac{\rho'\eta'' - \rho''\eta'}{((\rho')^2 + (\eta')^2)^{\frac{3}{2}}}, \tag{5.21}$$

where $\rho = log\left(\left\| g * \widehat{p_s}(\lambda) - p_m \right\|^2\right)$, $\eta = log(\|\widehat{p_s}(\lambda)\|^2)$ and superscript $'$ symbolizes the derivative with respect to $\lambda$. The solution to Eq. (5.21) is obtained numerically with Hansen's Regularization Tools [65].

### 5.6.2 **Generalized cross-validation**

The cross-validation method consists in calculating the reconstruction after removing a measurement point from the data, to compare the reconstruction with the actual data at that point, and to minimize the norm of this difference. This is performed for each data point, and to avoid costly computation, a generalized cross-validation (GCV) formula was derived analytically. This method can be used to determine the regularization parameter in Tikhonov's method by minimizing the cost function [29]:

$$\mathrm{J}_{GCV} = \frac{\|G \cdot \widehat{p_s} - p_m\|^2}{\left(Tr\left(I - \frac{|G|^2}{|G|^2 + \lambda}\right)\right)^2}, \tag{5.22}$$

with $Tr()$ symbolizing the trace and $G$ being the 3-D convolution matrix. When the operator is circulant or block-circulant, the elements on the main diagonal are constant and Eq. (5.22) becomes [29]:

$$\mathrm{J}_{GCV} = \frac{\sum \frac{\lambda}{\lambda + |\mathcal{F}\{g\}|^2} \cdot |\mathcal{F}\{p_m\}|^2}{\left(\sum \left(\frac{\lambda}{\lambda + |\mathcal{F}\{g\}|^2}\right)\right)^2}. \tag{5.23}$$

The summations are carried out over all elements in the three dimensions. Note that Eq. (5.23) does not require explicit computation of the convolution matrix $G$.

### 5.6.3 **Bayesian approach**

The idea behind the Bayesian method is to use Bayes' theorem to relate the measured pressure and the source field, which are both considered probabilistic functions. Prior knowledge of the source is included in the model to obtain a regularization cost function and corresponding filter. By considering Gaussian probability density functions (the most conservative assumption) for both the likelihood $P(p_m|p_s)$ and prior $P(p_s)$, a regularization cost function equivalent to Tikhonov's formula is obtained by determining the maximum a priori. In this form, the regularization parameter is interpreted as the noise-to-signal ratio.

Pereira, Antoni and Leclère showed that the Bayes' theorem approach can the applied to obtain a relation between the evidence $P(p_m)$, the prior and the likelihood with respect to the signal to noise ratio. Considering the mathematical assumptions described in [30], such relationship is used to derive the Bayesian the cost function expressed in Eq. (5.24). Its minimization predicts the optimal regularization parameter.

$$\mathrm{J}_{Bayes} = \sum ln(|\mathcal{F}\{g\}|^2 + \lambda) + N ln\left(\frac{1}{N}\sum \frac{|\mathcal{F}\{p_m\}|^2}{|\mathcal{F}\{g\}|^2 + \lambda}\right), \tag{5.24}$$

with $N = N_x \cdot N_y \cdot N_t$. The summations are carried out over all elements in the three dimensions.

### 5.6.4 **Comparison of the methods**

The circular baffled piston model described in Appendix A is used to generate the measured pressure field $p_m$. Four different levels of white Gaussian noise are added to $p_m$ to simulate experimental noise. Then, the regularized reconstruction is computed for regularization parameters ranging from $\lambda = 10^{-6}$ to $\lambda = 10^{1}$. The relative RMS errors obtained are presented by the black curves in Figure 5.8 for back-propagation distances of $d = 2$ cm and $d = 18$ cm. The square array

has a length of 2.4 m with spatial discretization of $dx = 2$ cm and the time domain discretization is $dt = 0.02$ ms over a period of 5 ms. In each case, the predicted optimal regularization parameter is calculated for the three studied methods by minimizing Eqs. (5.21), (5.23) and (5.24), and the results are shown by the markers in Figure 5.8. The color of the marker is associated to the signal-to-noise ratio (SNR) and the shape represents the prediction method.

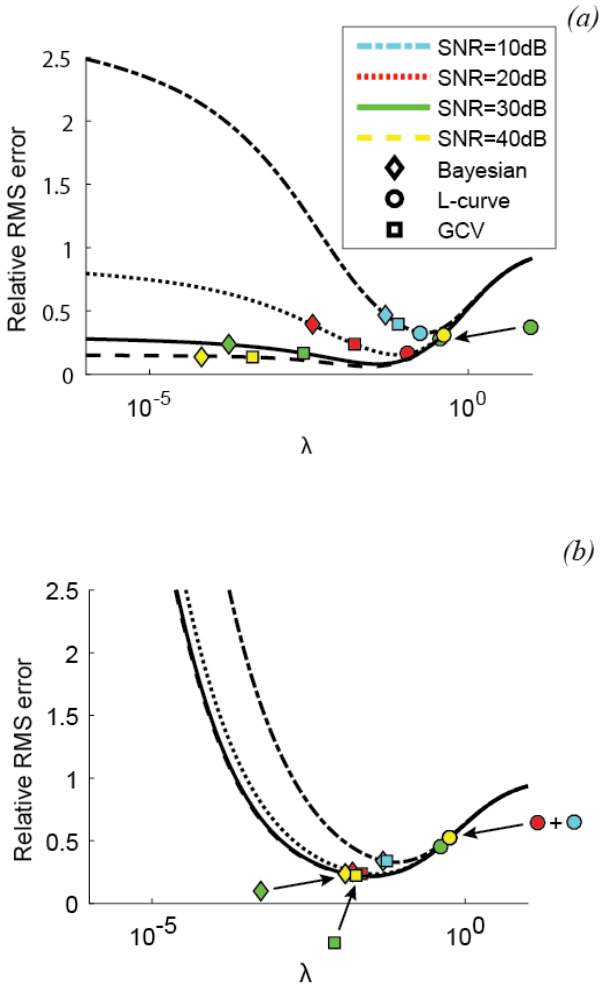


*Figure 5.8 : Relative RMS error with respect to the regularization parameter for different SNR values and propagation distances. (a) $d = 2$ cm, (b) $d = 18$ cm*

We observe that the steepness of the relative error curves increases with the noise level and with the back-propagation distance. For example, in Figure 5.8(a), the minimum error for the SNR of 10 dB is well-defined in comparison to that of the SNR of 40 dB for which the curve is shallower. For the 40 dB case, all regularization parameters values below $10^{-1}$ lead to similar relative errors.

The predictions of the regularization parameters are in general more accurate when the ill-posedness of the problem is more important. For lower noise level and propagation distances, all three regularization coefficients predicted are relatively far from the optimum. For greater propagation distances (Fig. 9(b)), the three methods yield coefficient values that are closer to the optimum.

In Figure 5.9, we present the difference between the relative RMS errors obtained with the predicted values of $\lambda$ and the minimum RMS error obtainable with the Tikhonov regularisation. These results are presented with respect to the back-propagation distance and signal-to-noise level.

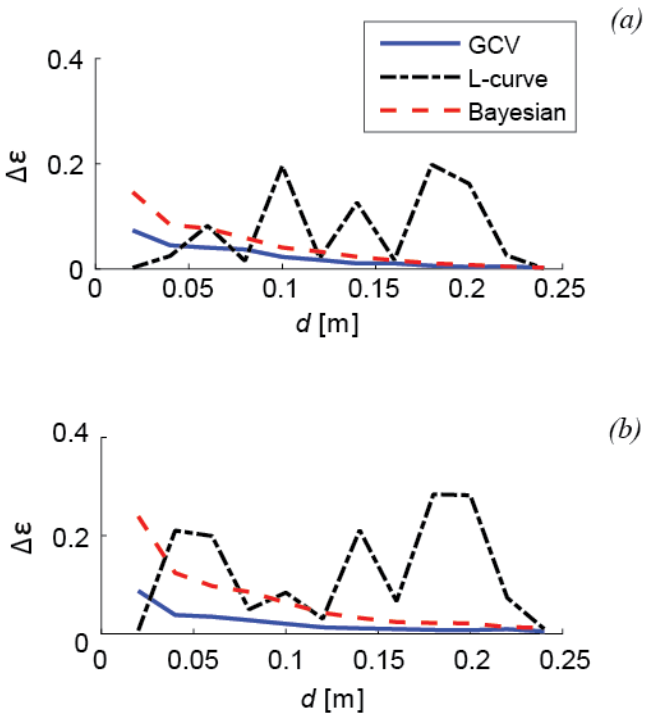


*Figure 5.9 : Difference of relative RMS error between the predictions of the GCV, the L-curve and the Bayesian cost functions with respect to the back-propagation distance. (a) SNR of 10 dB. (b) SNR of 20 dB. (c) SNR of 30 dB. (d) SNR of 40 dB*

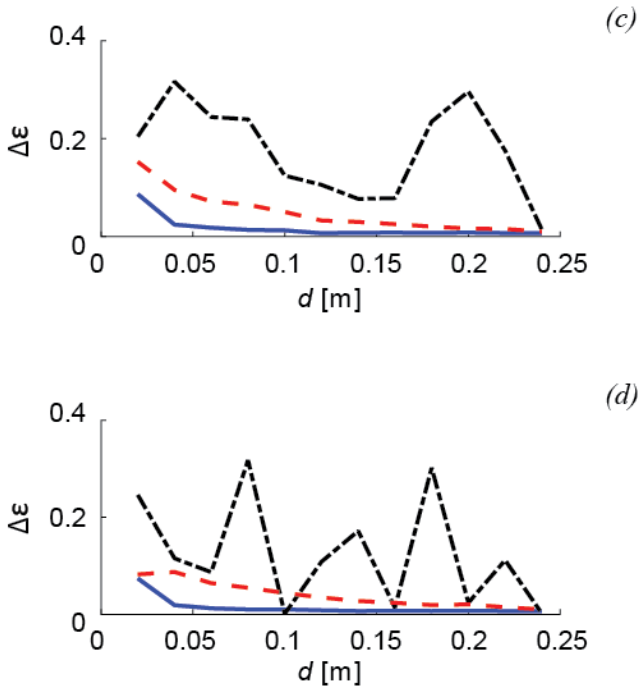


*Figure 5.9 (suite) : Difference of relative RMS error between the predictions of the GCV, the L-curve and the Bayesian cost functions with respect to the back-propagation distance. (a) SNR of 10 dB. (b) SNR of 20 dB. (c) SNR of 30 dB. (d) SNR of 40 dB*

For all cases studied, both the GCV and the Bayesian approach give satisfactory results. The GCV slightly outperforms the Bayesian approach. The predictions are more accurate for larger values of $d$, as the steepness of the prediction method's cost functions increases with the ill-posedness of the problem, as observed in Figure 5.8.

In general, the L-curve gives inconsistent results, even though it outperforms the other two methods in a very few cases. One explanation behind the poor performance of the L-curve is the fact that in some cases, the L-curve does not have an apparent negative curvature; in other words, the L-curve obtained has the shape of a "reversed" L. In other cases, the maximum curvature of the L-curve does not correspond to the optimal regularization parameter. We show such example in Figure 5.10.

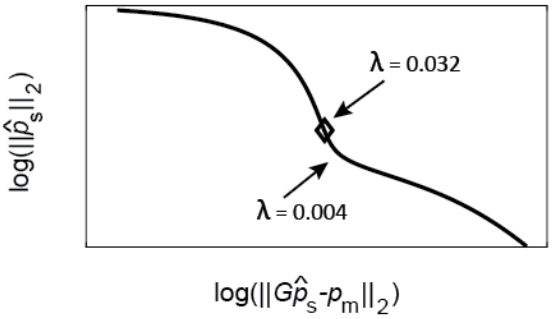


*Figure 5.10 : L-curve for $d = 20$ cm and SNR = 40 dB*

The diamond marker shows the position of the regularization coefficient that minimizes the RMS error. In this example, the maximum curvature corresponds to a regularization parameter of $\lambda = 0.004$ which is almost ten times smaller than the optimal value. The relative error obtained with $\lambda = 0.004$ is more than twice of that obtained with $\lambda = 0.032$.

The conclusion of this comparative analysis does not correspond to that of Williams [56] and Pereira et al. [30]: in their investigations, which were applied to stationary signals, the GCV did not give very good predictions. Perhaps the sparse nature of the studied non-stationary signals favors the performance of the GCV.

## 5.7 Conclusion

In this paper, we introduced a method for reconstructing non-stationary acoustic fields based on the 3-D linear deconvolution and the TSD Green's function. We compared the proposed formulation to the Fourier-based operator, and we have shown that the amplitude and the shape of the reconstruction is improved significantly with the proposed method. For typical back-propagation nearfield distances, that is up to a dozen centimeters, the RMS error obtained with the proposed method is as much as 62 % smaller than that of the standard method. With the proposed formulation, linear deconvolution is obtained by embedding the convolution matrix is a bigger block-circulant matrix. The process introduces truncations errors if the field is not fully sampled. We have shown that it is not possible to sample the entire field in both space and time; however, truncation error can be reduced by applying the patch extrapolation algorithm. We showed that its convergence can be obtained with the proposed method if a proper regularization parameter is used.

We compared three methods for predicting the optimal regularization parameter, and it was shown that the generalized cross-validation outperforms the L-curve and the empirical Bayesian approaches with the TSD method, for non-stationary signals studied. Future work could include integration of the 3-D linear deconvolution operator to the equivalent source method [41, 42] or the statistically optimized NAH [64], which are less restrictive in terms of the problem's geometry and number of sources. In addition, it would be interesting to investigate other regularization cost functions, perhaps one that takes advantage of the sparse nature of the non-stationary signals studied.

## 5.8 Appendix A. Circular baffled transient piston model

An analytical model to determine the transient radiation of a circular baffled piston was developed by Stepanishen [45]. The model is based on Green's function formulation and is exact in both the near-field and the far-field. It is applicable to any displacement of the piston. Since the displacement of the piston is uniform in space, Stepanishen shows that the pressure field with respect to time and position $r$ is expressed as:

$$p(r,t) = \rho_0 h(r,t) * \frac{\partial^2 x(t)}{\partial t^2}. \tag{5.25}$$

The $*$ symbol denotes the time domain convolution of the acceleration with the impulse response $h(r,t)$, and $\rho_0$ is the density of air. When piston radius $a > r$, the impulse response function is defined as:

$$h(r,t) = \begin{cases} 0, & ct < l \\ c, & l < ct < R_{min} \\ \frac{c}{\pi} cos^{-1}\left(\frac{(ct)^2 - l^2 + r^2 - a^2}{2r\sqrt{(ct)^2 - l^2}}\right), & R_{min} < ct < R_{max} \\ 0, & ct > R_{max} \end{cases} \tag{5.26}$$

When $a < r$, the impulse response function becomes:

$$h(r,t) = \tag{5.27}$$

$$\begin{matrix} 0, & ct < R_{min} \\ \frac{c}{\pi} cos^{-1}\left(\frac{(ct)^2 - l^2 + r^2 - a^2}{2r\sqrt{(ct)^2 - l^2}}\right), & R_{min} < ct < R_{max} \\ 0, & ct > R_{max} \end{matrix}$$

where $R_{min}$ and $R_{max}$ are respectively the shortest and longest distances from the observation point to the circumference of the piston:

$$\begin{aligned} R_{min} &= \sqrt{z^2 + (a-r)^2}, \\ R_{max} &= \sqrt{z^2 + (a+r)^2}, \end{aligned} \tag{5.28}$$

and $l$ is the projection of $r$ over the $z$ axis. In the analysis, an under-damped sinusoidal piston motion profile is used:

$$x(t) = \mathrm{A} \cdot \sin(2\pi f t) \cdot e^{-\zeta t}, \tag{5.29}$$

$$\frac{\partial^2 x(t)}{\partial t^2} = Ae^{-\zeta t}\left(-\omega^2 sin(\omega t - \pi/2) - 2\zeta\omega cos(\omega t - \pi/2) + \zeta^2 sin(\omega t - \pi/2)\right), \tag{5.30}$$

where $\zeta$ is the decay coefficient. The physical parameters involved in the calculation are presented in Tableau 5.1.

*Tableau 5.1: Physical parameters of the transient piston*

| *Parameters* | *Values* |
|---|---|
| Piston radius | $a = 0.20$ m |
| Sinusoidal frequency | $f = 500$ Hz |
| Decay coefficient | $\zeta = 1000\text{s}^{-1}$ |
| Amplitude | $A = 0.0001$m |
| Speed of sound | $c = 343$ m/s |
| Air density | $\rho_0 = 1.2$ kg/m$^3$ |
| Distance between the reconstruction and source planes | $z_0 - d = 0.005$ m |

# CHAPITRE 6 ARTICLE 3 : SPARSE REGULARIZATION FOR TIME DOMAIN NEARFIELD ACOUSTICAL HOLOGRAPHY

J.-M. Attendu, A. Ross, Soumis au *Journal of the Acoustical Society of America* le 21 novembre 2017.

## 6.1 Préface

Cet article s'intéresse à l'application de la régularisation $\ell^1$ à la reconstruction de signaux causaux, conformément à l'objectif *iii*. La régularisation $\ell^1$ favorise la convergence de la solution pour des valeurs proches de zéro : elle s'applique donc aux cas pour lesquels la solution attendue contient une grande proportion de valeurs nulles. Avec des sources non-stationnaires causales, cette méthode permet de réduire l'erreur causale, c'est-à-dire l'erreur qui survient dans le régime pour lequel le champ est nul, donc avant l'arrivée du premier front d'onde à une position et un temps donné. De même qu'avec la régularisation de Tikhonov, cette méthode nécessite de déterminer un paramètre de régularisation : on démontre que la courbe de Pareto produit une prédiction adéquate du paramètre pour les cas étudiés, ce qui permet d'atteindre l'objectif *iv*. On valide la méthode avec des signaux simulés et expérimentaux. Un désavantage de la régularisation $\ell^1$ par rapport à la régularisation $\ell^2$ est qu'elle nécessite de résoudre numériquement un problème convexe pour un grand nombre de données. La FFT est donc utilisée afin de limiter les temps de calcul, ce qui respecte l'objectif *v*.

## 6.2 Abstract

In this paper, we apply the $\ell^1$-norm sparse regularization method to time domain reconstruction of causal non-stationary acoustic fields such as impulse noise. This method properly reconstructs the back-propagated sound field where its amplitude should be null: for causal sources, this occurs mostly for positions and time that precede the arrival of the first wave front. Therefore, it suppresses causal errors typically found in time domain reconstruction when standard Tikhonov regularization is applied. We compare the reconstructions obtained from both Tikhonov and sparse regularization methods using a transient baffled piston model, and show that the global RMS error is significantly reduced when using sparse regularization. The improvement provided depends on the level of

sparsity of the reconstructed signal. For the studied cases, it can represent a reduction of the global root mean square error by up to a factor three. The performance of Pareto frontier curve for predicting the optimal sparse regularization parameter is examined; it leads to accurate predictions especially for lower noise levels. Finally, sparse regularization is applied to experimental data in order to obtain an accurate reconstruction of the sound field produced by an impacted plate.

## 6.3 Introduction

Nearfield acoustical holography (NAH) is the state-of-the-art method for visualizing acoustic fields over the three-dimensional space [1, 3, 2]. It requires measuring the sound field with a 2-D microphone array conformal to the source and located in its nearfield. This measurement acts as a boundary condition used to determine the sound field towards the source by means of a deconvolution with a Green's function derived from the wave equation.

When the studied source is not stationary, the sound field is represented in the time domain instead of the frequency domain, since the frequency components of the measured signal are not statistically constant over time. A method based on the use of the three-dimensional linear convolution and the time and spatial domains (TSD) Green's function was developed to accurately visualize non-stationary fields, for both the direct [54] and inverse [66] problems. An asset of this method is that the 3-D linear convolution completely suppresses wrap-around errors and the use of the TSD Green's function significantly reduces spectral leakage when the sound field is represented in the TSD.

It is well known that the inverse problem is ill-posed; it requires regularization since high frequency components with low signal to noise ratio (SNR) are amplified in the reconstruction process [9]. In NAH, this amplification is considerable since the measurement performed in the nearfield takes into account the contribution of evanescent wave components, which are amplified exponentially in the back-propagation process. The standard regularization method for NAH is Tikhonov regularization [7], however, such method fails at properly reconstructing the sound field where its amplitude is null. The main idea behind this work is that sparse regularization, also known as least absolute shrinkage and selection operator (LASSO) in statistics [23], which minimizes the $\ell^1$-norm of the solution instead of the $\ell^2$-norm, is better suited for the regularization of non-stationary acoustic fields. This assumption is based on the fact that this regularization method favors the

convergence of the solution to null values, which constitute a significant proportion of the causal non-stationary signals studied. On the contrary, Tikhonov regularization tends to spread the energy of the reconstruction, rather than its sparsity.

To the author's knowledge, sparse regularization has not been applied to non-stationary NAH before. However, the idea has been explored in the context of stationary NAH. Chardon et al. have applied compressed sensing with sparse regularization to stationary NAH [43]. They have shown that sparse regularization acts as an alternative regularization technique, and that reconstruction can be achieved with a much lower sampling rate (sub Nyquist) by using compressed sensing. Regularization errors they obtained are of the same order of magnitude as those from Tikhonov's method. Fernandez-Grande et al. demonstrated the application of compressed sensing and sparse regularization to the equivalent source method (ESM), a variant of standard NAH. Sparse regularization is used to successfully obtain the reconstruction, and the resulting error is under that of Tikhonov regularization [44]. Sparse regularization method has also been applied to other methods related to NAH: Simard and Antoni applied $\ell^1$ regularization to acoustic source identification method [67], and Alqadah and Valvidia applied sparse regularization to nearfield electromagnetic holography with the ESM [68].

In all of these cases, the reconstruction is represented in the frequency domain: the analysis deals with signals that are sparse in the sense that their amplitude is spatially localized. This condition is rarely obtained for each frequency component, and consequently, the performance of the reconstruction varies with the frequency. For the non-stationary case, the representation of the sparsity is different. We suppose that the studied signals are of finite duration with null amplitude before $t = t_0$ and after its duration $\tau + t_0$. The resulting sound field radiates from the source, and according to the nature of the studied sources, the field is null before its first wave front reaches the positon at which the reconstruction is performed. Sparse regularization properly reconstructs these null components and therefore suppresses causal errors obtained when Tikhonov regularization is applied.

The main objective of this paper is to compare regularization error obtained with the standard Tikhonov method with that of sparse regularization when applied to reconstructing causal sources. The comparison is performed using sound field obtained from a transient baffled piston model, and the analysis is developed for planar geometry, but could be extended to other geometries in future

work. We also investigate the performance of the Pareto frontier curve [26] (an adaptation of the L-curve method [25]) to predict the optimal regularization parameter for sparse regularization. In previous work, the method was successfully applied to electrical impedance tomography [27]. To the authors' knowledge, a method for predicting the optimal regularization parameter has not been proposed yet in the context of NAH. Finally, sparse regularization method is applied to experimental measurements obtained from non-stationary acoustic field radiating from an impacted plate.

## 6.4 Overview of time domain NAH with 3-D linear deconvolution and regularization

In this section, we present an overview of 3-D linear deconvolution formulation for time domain NAH, along with the two regularization methods compared in this paper.

### 6.4.1 Time domain NAH with the TSD operator

The geometry of the problem is illustrated in Figure 6.1. The finite acoustic source represented in gray is located on the source plane at $z = 0$. For the direct problem, the measurement is performed in the nearfield of the source at $z = z_0$, and the field is determined on plane $z = z_0 + d$, with $d$ being the propagation distance. The mathematical relationship between the pressure field $p$ on the two planes is a 3-D convolution, as expressed in Eq. (5.1):

$$p(t, x, y, z_0 + d) = \iiint_{-\infty}^{\infty} p(t', x', y', z_0) \cdot g(t - t', x - x', y - y', d) dt' dx' dy', \tag{6.1}$$

where $g$ is the time and spatial domain Green's function expressed as:

$$g(t, x, y, d) = \frac{d}{2\pi c}\left(\frac{c}{R^3} + \frac{1}{R^2}\frac{\partial}{\partial t}\right)\delta(t - R/c), \tag{6.2}$$

with sound velocity $c$ and propagation radius $R = \sqrt{x^2 + y^2 + d^2}$.

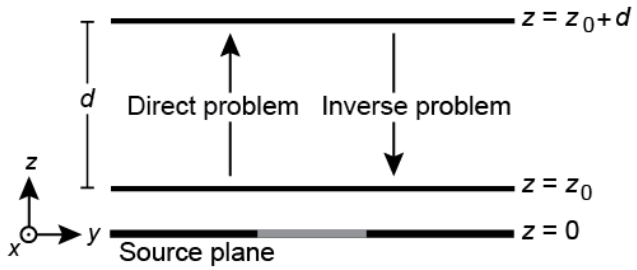


*Figure 6.1 : Geometry of planar NAH for direct and inverse problems*

The inverse problem, that is to obtained pressure $p(z_0)$ from a measurement on plane $p(z_0 + d)$, is obtained by discretizing Eq. (5.1), and expressing the convolution in the form of a matrix product:

$$p_m = G \cdot p_s. \tag{6.3}$$

Here, matrix $G$ takes the form of a Toeplitz block-Toeplitz convolution matrix [10], $p_s$ is the unknown sound field near the source, and $p_m$ is the measurement. The reconstruction $p_s$ is obtained by multiplying $p_m$ with the inverse of $G$. Matrix $G$, in its Toeplitz form, typically contains too many elements for its inverse to be computed with a standard desktop computer. As detailed in [66], an approximation of its inverse is obtained by embedding the Green's function in a larger block circulant matrix by zero-padding $g$; this requires the Green's function to be expressed in the TSD. Such block circulant matrix is diagonalized by the 3-D Fourier operator [19], and the calculation of its inverse can be performed efficiently with the fast Fourier transform (FFT):

$$\widetilde{p_s} \approx \mathcal{F}^{-1}\{\mathcal{F}\{p_m\}/\mathcal{F}\{\tilde{g}\}\}, \tag{6.4}$$

where the tilde represents application of 3-D zero-padding on the signal and $\mathcal{F}$ and $\mathcal{F}^{-1}$ are the direct and inverse 3-D FFT operators.

### 6.4.2 **Tikhonov regularization**

Tikhonov regularization consists in finding a compromise between the residual norm $\|Gp_s - p_m\|_2$ and the norm of the solution $\|\Gamma p_s\|_2$, where $\|\cdot\|_2$ denotes the $\ell^2$-norm. In this paper, we assume

the Tikhonov matrix Γ to be the identity, and the solution to this regularization problem becomes analogous to applying a low-pass filter. When the amount of regularization applied is not significant (corresponding to a high cut-off frequency), the amplified noised components dominate the reconstruction and the norm of the solution $\|p_s\|_2$ becomes large. At the opposite, when too much regularization is applied, $\|p_s\|_2$ tends to zero and the residual norm $\|Gp_s - p_m\|_2$ becomes large. The optimal solution lies in-between these two cases, and such compromise if found by minimizing $p_s$ in the Lagrangian expression:

$$argmin\left(\left\|G\widetilde{p_s} - p_m\right\|_2^2 + \lambda \cdot \left\|\widetilde{p_s}\right\|_2^2\right), \tag{6.5}$$

where the regularization coefficient $\lambda$ acts as a Lagrange multiplier. The solution $\widehat{p_s}$ is obtained by taking the derivative of (5.4) and equalizing to zero:

$$\widehat{p_s} = (\mathrm{G}^{\mathrm{T}}G + \lambda)^{-1}(G^T p_m). \tag{6.6}$$

In the Fourier domain, regularization term in Eq. (6.6) is in fact the low-pass Wiener filter expressed in Eq. (6.7), and the regularization parameter $\lambda$ acts as its cut-off frequency:

$$\mathrm{W} = \frac{|\mathcal{F}\{\tilde{g}\}|^2}{|\mathcal{F}\{\tilde{g}\}|^2 + \lambda}. \tag{6.7}$$

The operations required for solving the inverse problem with Tikhonov regularization are described in the block diagram below:

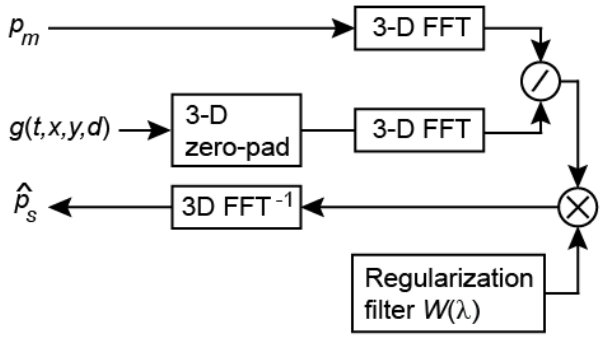


*Figure 6.2 : Block diagram of the operations involved in the 3-D linear deconvolution with Tikhonov regularization*

The first step is to apply 3-D zero-padding to the TSD Green's function in order to double its values in the time and spatial domains. Padding can also be applied to the measured field $p_m$ in order to fit the size of the Green's function if needed. Then, a three dimensional FFT is applied to both signals, and the measured field is divided by the Green's function in the Fourier domain. The Wiener filter with parameter $\lambda$ is applied to regularize the reconstruction. Finally, the inverse 3-D FFT is applied and the reconstruction $\widehat{p_s}$ is obtained.

Tikhonov regularization is not effective in reconstructing null components. This is due to the application of the $\ell^2$-norm to the solution $p_s$ in Eq. (5.4): by evaluating the null values of the solution $p_s$ as a non-zero values smaller than one, the application of the power of two on such values reduces their amplitudes and brings them close to zero. A consequence of this is the generation of ripples where the reconstructed pressure field should be null. In addition to creating causality errors, this considerably increases the global RMS error of the reconstruction.

### 6.4.3 **Sparse regularization**

The sparse regularization consists in exchanging the $\ell^2$-norm applied to the solution $p_s$ with the $\ell^1$-norm. The $\ell^2$ residual norm is conserved. This solves the problem described above since it gives a more significant impact in the cost function to the ripples where the solution should be null. The sparse regularization cost function is expressed as:

$$argmin\left(\left\|Gp_s - p_m\right\|_2^2 + \lambda \cdot \|p_s\|_1\right). \tag{6.8}$$

In this case, it is not possible to derive a solution in a closed form, and the solution $\widehat{p_s}$ to Eq. (6.8) must be computed numerically. Such problem is known as basis pursuit denoising: it is a convex optimization problem for which many algorithms were developed [24]. In this paper, the YALL1 module based on alternating direction algorithm is used on MATLAB [69]. An asset of this module is that the convolution matrix $G$ does not have to be explicitly expressed, as YALL1 can take advantage of FFT-based computation for the matrix multiplication. It was shown that both the accuracy and computational efficiency of the YALL1 module is competitive in comparison to other MATLAB based convex optimization algorithm applied to $\ell^1$ regularization [70]. The operations involved in solving the inverse problem with sparse regularization are described in the block diagram below:

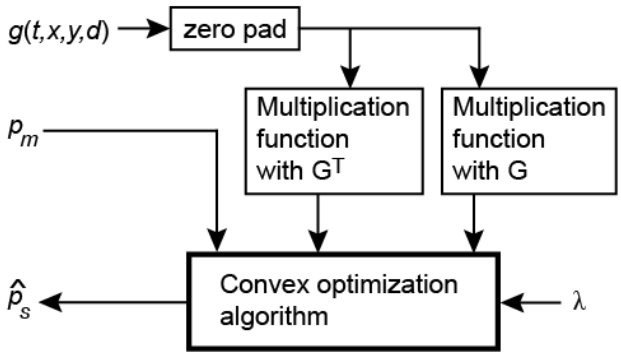


*Figure 6.3 : Block diagram of the operations involved in the 3-D linear deconvolution with sparse regularization*

The 3-D zero-padding is applied to the TSD Green's function to obtain the linear deconvolution formulation. It can also be applied to the measurement $p_m$ to fit the size of $g$. The Green's function is inputted into the convex optimization algorithm through two function handles: one for its multiplication, and another for the multiplication with its transpose. Matrix multiplication is performed with the FFT, taking advantage of the circulant nature of $G$ and making the operation computationally more efficient. The measured field and the regularization parameter are other inputs of the convex optimization algorithm.

### 6.4.4 **Bayesian interpretation**

Another way of interpreting the difference between Tikhonov and sparse regularization is through the Bayesian approach. We consider the measured pressure field $p_m$ and the pressure at the source $p_s$ as probabilistic quantities. One can be related to the other using Bayes' theorem:

$$\mathrm{P}(p_s|p_m) = \frac{\mathrm{P}(p_m|p_s) \cdot P(p_s)}{P(p_m)}. \tag{6.9}$$

The posterior $\mathrm{P}(p_s|p_m)$ is the probability density function (PDF) that expresses the probability of obtaining such pressure field at the source knowing the measurement $p_m$. By maximizing the posterior, that is to select the solution which has the most probable occurrence, the reconstruction can be determined. The posterior depends on the PDF attributed to the likelihood function

P$(p_m|p_s)$ and the prior $P(p_s)$. $P(p_m)$ acts as a normalization constant since the integral of the PDF must be unitary.

According to the central limit theorem, it makes sense to assign a Gaussian PDF to the likelihood function, since the variations of the measurement usually depends on many independent random processes, such as electronic or background noise. In addition, assuming Gaussian PDF is the most conservative choice, since it is the distribution with maximum entropy.

The prior $P(p_s)$ can take many forms, depending on our knowledge of the source. By assuming a Gaussian PDF, the maximum a posteriori becomes equivalent to the Tikhonov cost function expressed in Eq. (5.4), where $\lambda$ is interpreted as the noise to signal ratio (i.e. the ratio of the variance of the Gaussian prior and likelihood) [30]. For a more parsimonious source, a Laplace distribution can be assumed, which leads to the sparse regularization cost function Eq. (6.8). Gaussian and Laplace distributions with unitary variance $\sigma^2$ and mean $\mu$ are compared in Figure 6.4.

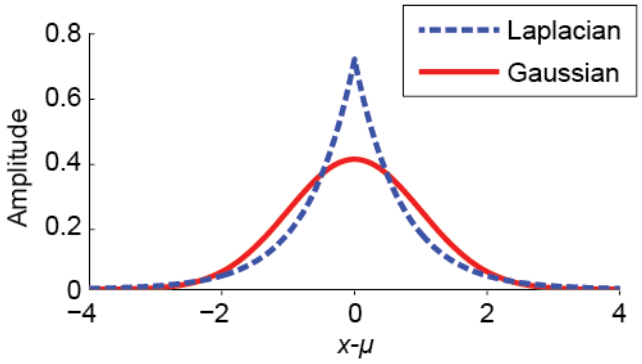


*Figure 6.4 : Comparison between the Gaussian and Laplacian probability density functions*

The Laplace distribution has a larger proportion of its amplitude contained near its maximum and its tails. This reflects its tendency to produce reconstruction values that are either relatively large or close to zero.

## 6.5 Comparative analysis

In this section, we present a comparative analysis of the regularization error obtained from the sparse and Tikhonov methods. The analysis is based on the use of a baffled circular piston model [45]. This theoretical model is exact in the nearfield for any transient displacement of the piston.

We consider an exponentially decaying sinusoidal displacement to simulate a damped vibrating source. An overview of the model and physical parameters used is presented in Appendix A.

### 6.5.1 Methodology

The approach is as follows:

i. A signal $p_s$ is generated on the reconstruction plane $z = z_0$ using the baffled circular piston model.

ii. The 3-D linear convolution of $p_s$ with the TSD Green's function is calculated to determine the propagated field on the plane $z = z_0 + d$.

iii. We add white Gaussian noise of given SNR to the propagated field, which acts as the measurement $p_m$.

iv. We solve the inverse problem with both regularization methods for a wide range of regularization parameters $\lambda$.

v. We calculate the relative RMS error between the reconstructions $\widehat{p_s}$ and the simulated pressure $p_s$ field for both methods.

An asset of this approach is that only the noise and its amplification due to the ill-posedness of the problem contribute to the error, since the operators for propagation and back-propagation are the exact inverse of each others. Therefore, other sources of error typically found in NAH, such as truncation error, do not contribute to the RMS error; the analysis focuses solely on the regularization problem, as intended.

At this point in the paper, we assume that we have means of predicting the optimal regularization parameter. Therefore, for each studied cases, we select the solution associated to the regularization parameter $\lambda_{opt}$ that minimizes the RMS error. The comparison is thus made between the optimal reconstructions for both methods.

### 6.5.2 Relative RMS error

The relative root mean square error expressed in Eq. (5.8) is used to compare the results obtained with the two methods.

$$\varepsilon = \frac{\sqrt{\sum_{xyt}(\widehat{p_s} - p_s)^2}}{\sqrt{\sum_{xyt} p_s^2}}. \tag{6.10}$$

Here, $p_s$ is the simulated sound field at $z = z_0$ and $\widehat{p_s}$ is the regularized reconstruction obtained from the noised measurement.

### 6.5.3 **Comparative example**

We first present in Figure  a comparative example of typical reconstruction results obtained with the two regularization methods. In this case, the distance between each microphone is $dx = 12$ cm, the microphone array span is 1.2 by 1.2 m wide and the back-propagation distance is $d = 6 \cdot dx$. White Gaussian noise is added to $p_m$ in order to obtain a SNR of 30 decibels.

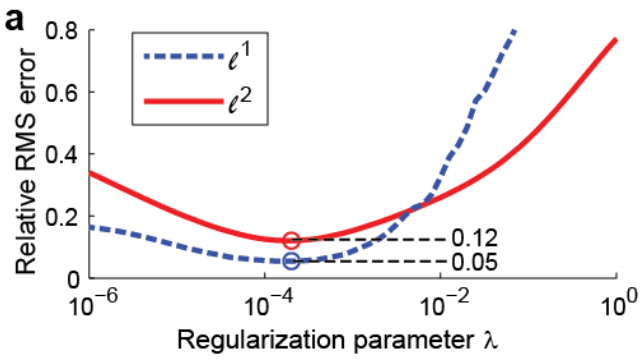


*Figure 6.5 : Example of reconstructions obtained with the Tikhonov ($\ell^2$) and sparse ($\ell^1$) methods. (a) Relative RMS error with respect to the regularization parameter $\lambda$. (b) Spatial reconstructions at $t = 2.4$ ms. (c) Absolute error of the spatial reconstruction at $t = 2.4$ ms. (d) Time domain reconstructions at $(x, y) = (0{,}0.36)$ m*

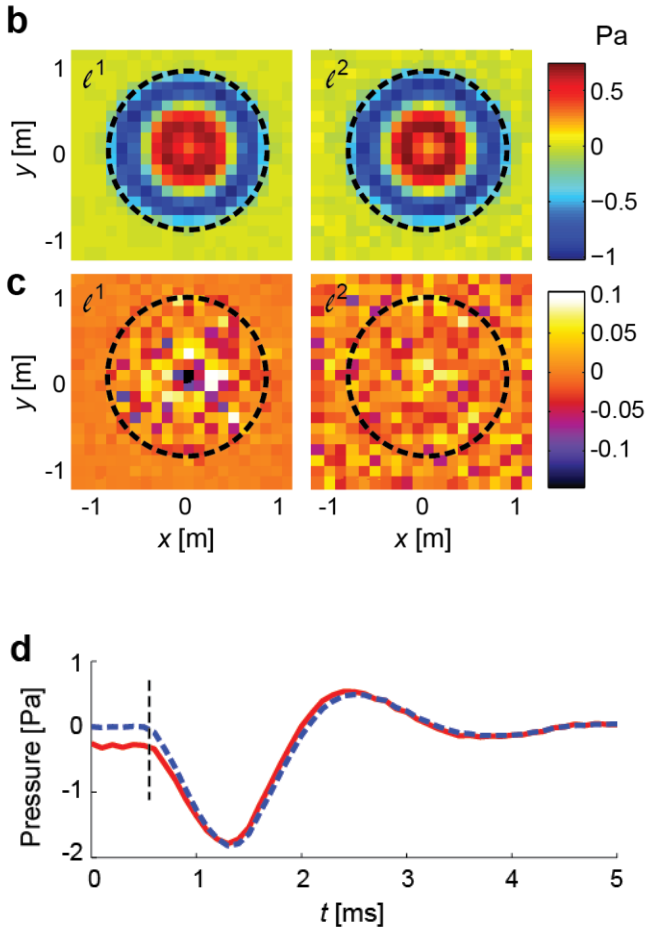


*Figure 6.5 (suite) : Example of reconstructions obtained with the Tikhonov ($\ell^2$) and sparse ($\ell^1$) methods. (a) Relative RMS error with respect to the regularization parameter $\lambda$. (b) Spatial reconstructions at $t = 2.4$ ms. (c) Absolute error of the spatial reconstruction at $t = 2.4$ ms. (d) Time domain reconstructions at $(x, y) = (0,0.36)$ m*

In Figure 6.5 (a), we present the relative RMS error with respect to the regularization parameter $\lambda$. Tikhonov regularization produces a minimal relative RMS error of 12 % as opposed to 5 % with sparse regularisation: the global error is thus less than half as small with sparse method. Furthermore, the error with sparse regularization remains lower than with Tikhonov regularization for a wide range of regularization parameters (up to $\lambda = 5 \cdot 10^{-3}$) in this example. Consequently, the improvement due to the sparse method remains effective even if optimal regularization parameter is not accurately predicted. The improvement caused by sparse regularization is due to the sparse nature of the solution, as it is reflected in Figure 6.5 (b) and (c).

In Figure 6.5 (b), the spatial reconstructions $\widehat{p_s}$ obtained with both regularization methods are presented for $t = 2.4$ ms. The corresponding absolute errors $(\widehat{p_s} - p_s)$ are presented in Figure 6.5 (c). The black dashed circles represent the limit of the propagation of the sound field at $t = 2.4$ ms; beyond this limit, the causal pressure field should be null. With Tikhonov regularization, we observe that the absolute error is distributed uniformly over the spatial domain; noisy values beyond the propagation limit produce causal errors in the reconstruction. With sparse regularization, the absolute error beyond the propagation limit drops. The average RMS error beyond the propagation boundary is more than four times smaller for sparse regularized reconstruction at 0.006 Pa than it is for Tikhonov regularized reconstruction at 0.028 Pa. However, within the boundary, the absolute RMS error is lower for Tikhonov regularization than it is for the sparse method (0.031 Pa vs 0.041 Pa). This reflects the fact that of the $\ell^1$-norm based method is effective only when the reconstructed signal has a high level of sparsity; otherwise, the standard Tikhonov method should be used.

In Figure 6.5 (d), the time domain reconstruction at $x = 0$ m and $y = 0.36$ m is presented. The black vertical dashed line represents the time it takes to the first wave front to travel from the piston and reach the position of observation. Again, before this limit, the amplitude of the pressure field should be null, which is approximately the case with the reconstruction obtained from the sparse regularization method. In fact, the reconstruction obtained with sparse regularization is almost identical to the solution.. With Tikhonov regularization, we notice the presence of causal errors. Beyond the limit, both methods lead to similar results.

### 6.5.4 **Analysis with respect to the back-propagation distance and noise level**

We present an analysis of the relative RMS error with respect to levels of white Gaussian noise and back-propagation distances. The relative RMS error is presented in Figure 6.6 (a). The relative improvement obtained from the difference in RMS errors between the two methods normalized by the relative RMS error of the sparse method is presented in Figure 6.6 (b). In each case, reconstruction with the regularization parameter which minimizes the error is selected. The area highlighted in orange in Figure 6.6 (b) represents the condition in which sparse regularization is outperformed by Tikhonov regularization.

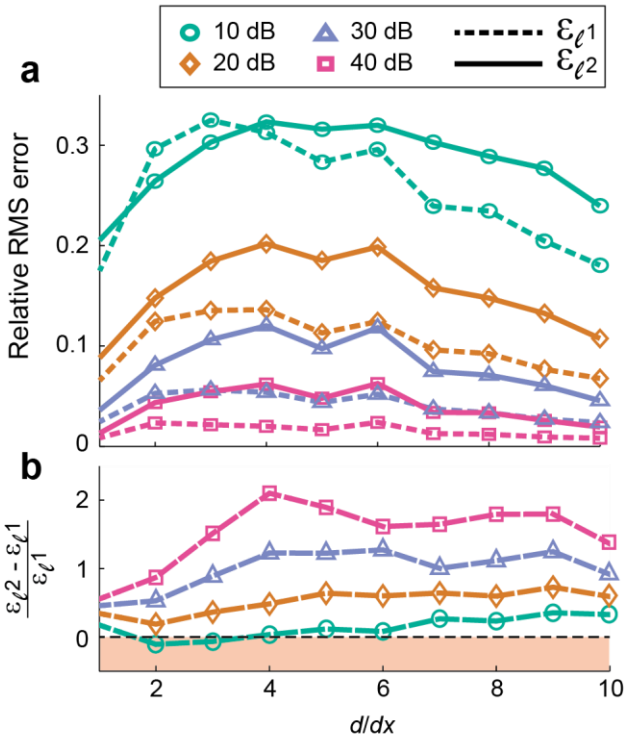


*Figure 6.6 : Error analysis of Tikhonov and sparse regularization with respect to the SNR and the normalized back-propagation distance. (a) Relative RMS errors ε. (b) Difference in RMS errors for both methods normalized by that of the sparse method*

We observe in Figure 6.6 (a) that the relative RMS error increases with the noise level, as expected. Additionally, the relative RMS error obtained with the sparse regularization method (dashed curves) is below that of the Tikhonov method (solid curves) for almost all cases studied.

For all SNR studied, the relative RMS error does not increase monotonically with the back-propagation distance: beyond about $d = 6 \cdot d/dx$, it decreases even though the ill-posedness becomes more important as $d$ increases. This is due to the fact that, for a given SNR, the amplitude of the measurement $p_m$ decreases with $d$, and consequently, the added noise level also decreases. Consequently, the absolute noise level added is more important for shorter propagation distances, and there is a trade-off between the added noise level and its amplification caused by the ill-posedness of the problem.

We observe in Figure 6.6 (b) that the improvement of the sparse method over the Tikhonov method decreases as the noise level increases, that is for lower SNR, and smaller back-propagation

distances. In fact, the improvement becomes inexistant for high noise levels because the parsimonious property of the signal is lost. Indeed, in the presented example, when the added noise is at its highest (10 dB and $d = \{2,3\} \cdot dx$), Tikhonov outperforms the sparse method.

### 6.5.5 Regularization error with respect to the sparsity of the measurement

To support the fact that the performance of sparse regularization increases with the degree of sparsity of the reconstruction, the error is analysed with respect to the span of the square array. In Figure 6.7, the proportion of sparsity increases with the size of the microphone array, because with wider spatial and time domain widows, it takes more time for the wave front to travel across the whole finite domain. In this case, the back-propagation distance is $6 \cdot dx$ with $dx = 12$ cm.

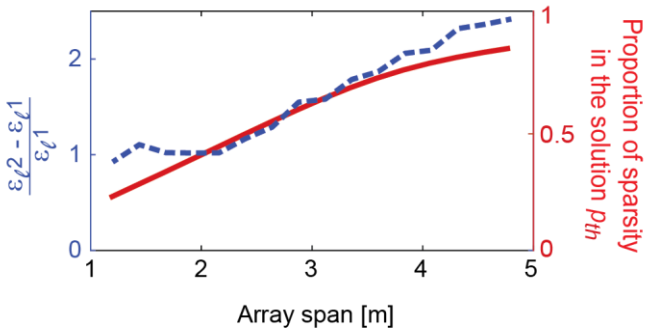


*Figure 6.7 : Sparsity ratio (solid curve) and difference in RMS errors for both studied methods normalized by that of the sparse method (dashed curve) with respect to the span of the square array*

In Figure 6.7, the proportion of sparsity is represented in red by the solid curve. It refers to the ratio of the number of zero elements to the number of all elements in the solution $p_s$. The improvement of sparse regularization with respect to Tikhonov regularization is presented by the blue dashed curve. Clearly, the improvement of sparse regularization with respect to Tikhonov method is due to the sparsity of the solution.

### 6.5.6 Computational efficiency

The main drawback of sparse regularization when compared to the standard method is its increased computational cost. Because the solution to $\ell^2$ regularization is obtained analytically, its computation is straightforward and very efficient; for a standard microphone array, it takes less

than a second on a desktop computer. The basis pursuit problem requires solving an iterative convex optimization algorithm. With the YALL1 module, for a tolerance of $10^{-4}$ and with $N$ in the order of $10^4$, the algorithm takes about 100 s to converge with MATLAB on a desktop computer. However, this calculation time could certainly be improved by lowering the floating-point precision or by using a more computationally efficient programming language.

## 6.6 Prediction of the optimal regularization parameter with the Pareto frontier curve

The accuracy of the sparse reconstruction algorithm depends on the choice of regularization parameter $\lambda$. Performance of standard methods for predicting $\lambda$ with sparse regularization such as the generalized cross validation, the Bayesian information criterion and the Akaike information criterion were investigated. However the results we obtained when applied in conjunction with the non-stationary NAH method proposed were of poor accuracy. Instead, in this section, we focus our analysis on the Pareto frontier curve, a variant of the L-curve method, which more adequately predicts the optimal regularization parameter. We first present the method, then, analysis of the predictions with respect to the SNR and the back-propagation distance is presented.

### 6.6.1 Pareto frontier curve

The Pareto frontier curve is analogous to the L-curve method frequently used in NAH for predicting the optimal regularization parameter with Tikhonov regularization. However, with sparse regularization, it considers the $\ell^1$-norm of the solution. Since Eq. (6.8) consists in a trade-off between the residual norm and the norm of the solution, the Pareto frontier curve involves plotting the logarithm of the two quantities for various values of $\lambda$, and to find a trade-off value where the curvature of the curve is maximal.

It is demonstrated in Ref. [26] that when applied to sparse regularization, such curve is convex and continuously differentiable over all points of interest, which ensures its applicability. An example of the Pareto frontier curve is shown in Figure 6.8 for $d = 5 \cdot dx$ and an SNR of 10 dB.

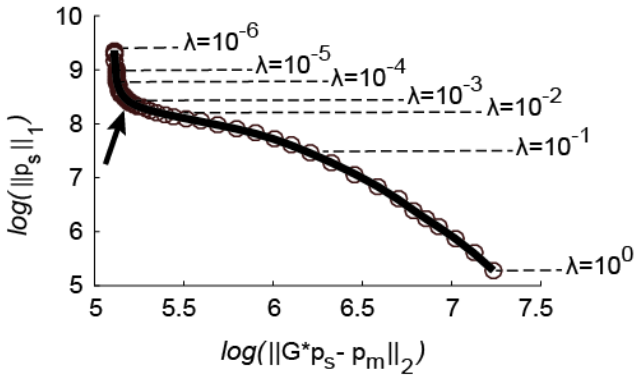


*Figure 6.8 : Pareto frontier curve for $d = 5 \cdot dx$ and a SNR of 10 dB*

The trade-off value for $\lambda$ is obtained at the maximum negative curvature located in the "elbow" of the L-shaped curve and denoted by the arrow in Figure 6.8. In this case, the prediction for the optimal value of $\lambda$ is about $10^{-3}$. The minimum curvature is obtained numerically by minimizing the function:

$$\mathrm{J}_{Lcurve} = \frac{\eta(\rho)''}{(1 + (\eta(\rho)')^2)^{\frac{3}{2}}}, \tag{6.11}$$

where $\rho = log\left(\left\| g * \widehat{p_s}(\lambda) - p_m \right\|_2\right)$ and $\eta = log(\|\widehat{p_s}(\lambda)\|_1)$. We calculated the derivative by interpolating with cubic spline functions.

Furthermore, Eq. (6.11) can be expressed as a function of $\lambda$, thus the minimum can conveniently be obtained numerically without computing the solution $p_s$ for numerous regularization parameters, therefore speeding up the optimization process [26].

### 6.6.2 **Prediction results**

In Figure 6.9, we present the relative RMS error with respect to the regularization parameter. The results are shown for three back-propagation distances and four noise levels. In each case, the black diamond markers indicate the predictions obtained with the Pareto frontier curve.

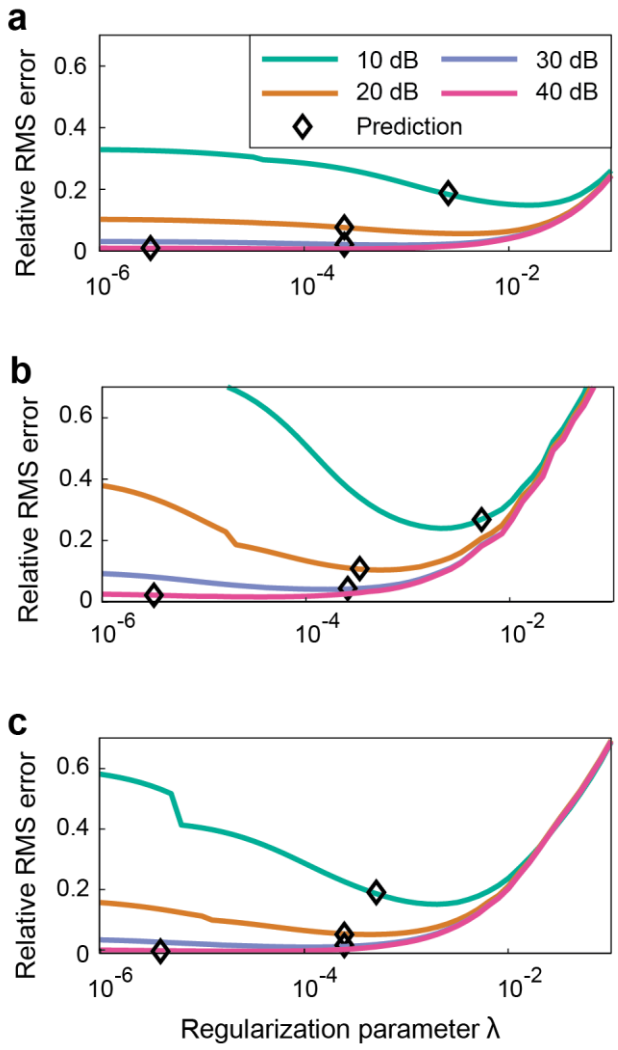


*Figure 6.9 : Relative error with respect to the regularization parameter. (a)* $d = 1 \cdot dx$*. (b)* $d = 5 \cdot dx$*. (c)* $d = 9 \cdot dx$

We observe in Figure 6.9 that for all back-propagation distances, at high SNRs (low noise levels), the relative RMS error does not vary significantly with the regularization parameter for low values of $\lambda$. For higher values of $\lambda$, the error starts increasing more significantly. Since the relative error curve remains flat with respect to $\lambda$, the accuracy of the reconstruction is not very sensitive to regularization coefficient, as long as the prediction remain under about a threshold of $\lambda = 10^{-3}$ for the examples presented. In all presented cases with high SNR, Pareto frontier curve method accurately predicts the minimum.

For lower SNR (high noise levels), the relative RMS error becomes more sensitive to the choice of the regularization parameter: the curves in Figure 6.9 for a SNR of 10 dB is much steeper, especially for $d = 5 \cdot d/dx$ where the error is high (see Figure 6.6 (a)). In these cases, prediction of the regularization parameter can have important effect on the accuracy of the reconstruction, especially since the RMS error is already high. For example, in Figure 6.9 (a), the minimum is at $\lambda = 1.6 \cdot 10^{-2}$ and the predicted value is $\lambda = 2.5 \cdot 10^{-3}$: this corresponds to increasing the error from 14.9 % to 18.4 % (about 20 % increase).

In Figure 6.10, we present the difference in relative RMS error between the optimal reconstruction and the reconstruction obtained with the regularization parameter predicted with the Pareto frontier curve method.

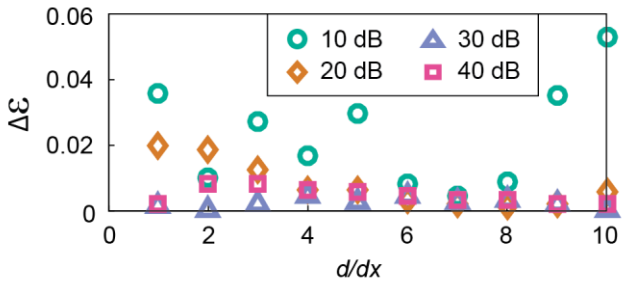


*Figure 6.10 : Difference in relative RMS error between the error obtained from the predicted regularization parameter and the minimum error with respect to the back-propagation distance*

In general, the Pareto frontier curve method produces a good prediction of the optimal regularization parameter. The difference in relative error remains low (under 1 % in most cases), especially for higher SNR. The difference becomes more important at lower SNR, however, since the absolute error is already important, the relative effect on the result is not significant.

The prediction error Δε remains adequate: by adding it to the optimal sparse regularization error in Figure 6.6 (a), the total error remains under that of the minimal error produced with Tikhonov method.

## 6.7 Experimental application

In this section, both sparse and Tikhonov regularizations are applied to the reconstruction of a transient field measured in an anechoic environment. The field is produced by the impact of a

metallic rod on a poly(methyl methacrylate) (PMMA) plate. The reconstructions obtained with both methods are compared.

### 6.7.1 Experimental setup

The experimental setup considered is presented in in Figure 6.11. The studied acoustic source is a 500 by 300 by 6 mm PMMA plate impacted at its center. The plate is represented by the grey surface in Figure 6.11. It is fixed to the frame of the apparatus using four metallic wires attached at its corners, which allows the plate to move freely in the vertical axis. The impact rod is released by an electromagnet, which ensures the repeatability of the impact conditions.

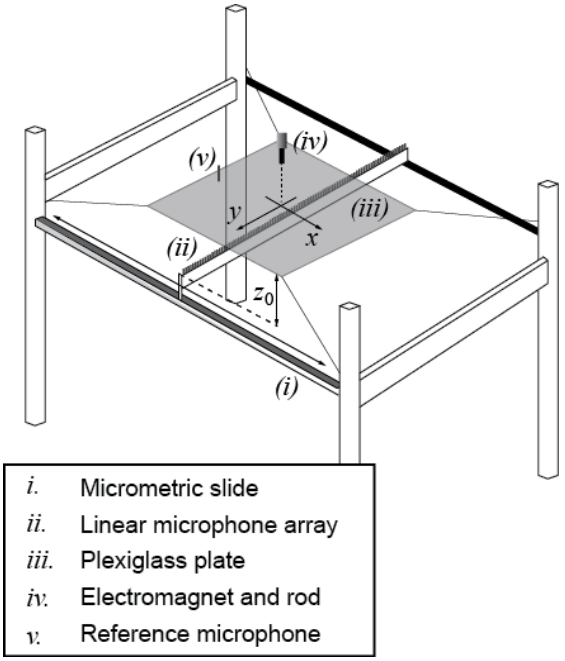


*Figure 6.11 : Experimental setup. Adapted from [54]*

After the impact, the plate vibrates for several milliseconds before its movement is completely damped. The measurement is performed using a straight antenna made of 64 class-1 microphones distanced by one cm center to center. A micrometric slide allows translating the antenna over the scanning surface (91 cm by 63 cm), with an increment of 1 cm. The impact is repeated for each measurement position by releasing the rod with the electro-magnet, and the measurements are phased by using a $65^{th}$ fixed reference microphone. The electro-magnet and rod setup allows a high

repeatability of the plate excitation. The sampling frequency is 50 kHz, 1000 samples are acquired, and the acquisition is performed with a 24-bit resolution.

With sparse regularization, the regularization parameter is chosen with the Pareto frontier curve method as described in Section 6.6. With the Tikhonov regularization, the parameter is predicted using the generalized cross-validation, as described in [66]. In both cases, the back-propagation distance is $d = 5 \cdot dx = 5$ cm, which corresponds to the location of the plane source.

### 6.7.2 Experimental results

For the experimental results, the true values of the field at the reconstruction plane are unknown. Consequently, it is not possible to calculate the RMS error associated to the reconstruction for this case and thus, the analysis presented in this section is qualitative. The reconstructions with both methods are presented in Figure 6.12 in the spatial domain (for $t = 3.5\ ms$) and in the time domain at the center of the aperture as well as at $x = y = 15$ cm from the center.

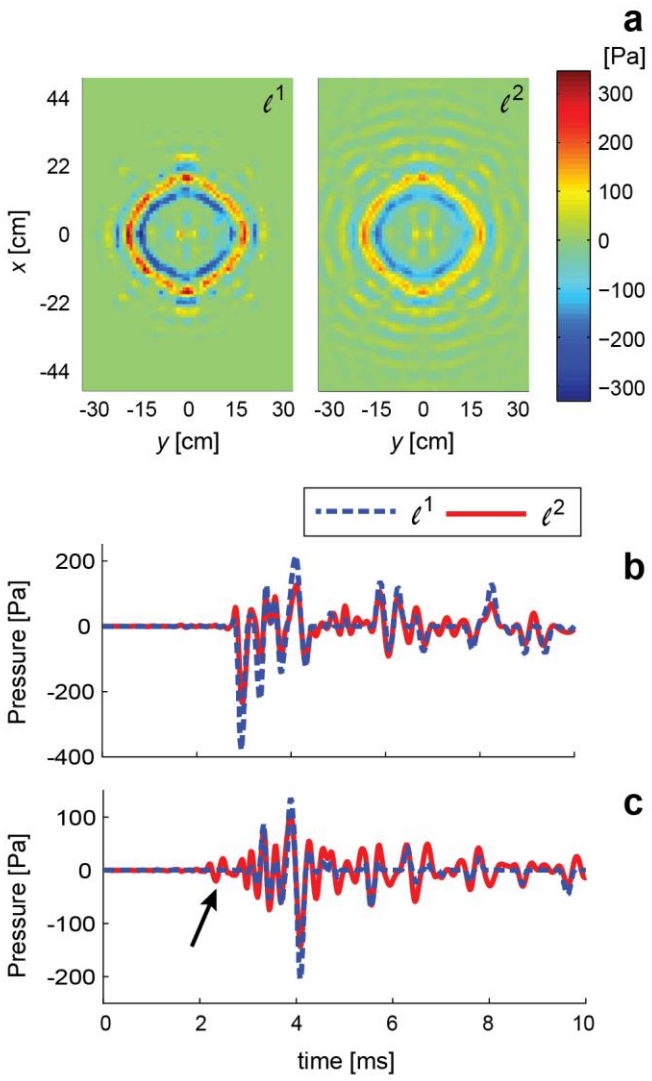


*Figure 6.12 : Reconstruction on the source plane obtained from experimental measurement with sparse and Tikhonov regularization methods. (a) In the spatial domain at time sample* $\boldsymbol{t = 3.5}$ *ms. (b) In the time domain at the center of the aperture* $(x, y) = (0,0)$ *m. (c) In the time domain at* $(x, y) = (0.15,0.15)$ *m*

When the center of the plate is struck by the rod, a flexural wave is generated in the plate and propagates radially with respect to the impact point. This is observable in the reconstructed acoustic field as shown in Figure 6.12 (a) in the spatial domain. The main wave front represented as a red circle corresponds to the primary rise in amplitude caused by the impact, and propagated over 3.5 ms. The wave front travels radially from its origin and is followed by a succession of oscillations (i.e. smaller concentric circles) with lower amplitude that corresponds to the damped displacement

of the plate that occurs following the impact. Because the flexural wave is dispersive, the presence of low amplitude, high frequency and high velocity wave components preceding the main wave front is also expected [71]. The larger concentric circles in Fig. 12(a) correspond to such high frequency components.

The presence of ripples beyond the main wave front is observed in the spatial domain for both regularization methods. With the $\ell^1$ regularization, the amplitude of the ripples falls down to zero a few centimeters beyond the boundaries of the main wave front. The selection of the regularization parameter with the sparse method seems adequate, since the amplitude of the field is not null around the main wave front, thus supporting the presence of high frequency components, while suppressing causal errors at positions close to the boundaries of the finite aperture. However, with $\ell^2$ regularization, the extent of these ripples is such that it reaches the boundaries of the finite aperture. A majority of these components should correspond to non-causal deconvolution errors, since the dispersive components travel only slightly faster than the main wave front, as their velocity is proportional to the square root of the band-limited frequency [13].

In the time domain, both methods lead to comparable results on the central position, as shown in Figure 6.12 (b). The maxima are somewhat higher with sparse regularisation and some low amplitude oscillations are attenuated to zero. This oscillatory behavior of the plate is expected, since the vibration field should not contain any discontinuity, and it is likely that the sparse method over attenuates these components. Similar conclusions are drawn for the time domain results obtained at $x = y = 0.15$ and presented in Figure 6.12 (c). However, in this case, the reconstruction obtained with the Tikhonov regularization results in important oscillations from $t = 2$ ms (highlighted by the black arrow), even though the causal field should be null for that time and position. Indeed, since the wave originates around t = 2.7 ms at the impact point (as seen in Fig. 12 (b)), and because some time is required for its travel, it is expected that the wave will reach the position in Figure 6.12 (c) later than t = 2.7 ms. With sparse regularization, these causal errors are suppressed. In conclusion, the observations described for the simulated examples presented in Section 6.5 remain applicable to the experimental results: in general, with sparse regularization, the amplitude of the field beyond the main wave front converges to zero. The reconstruction of the causal field is thus improved by sparse regularization, as the causality of the field is preserved. This stands even with measurement noise, sampling frequencies and finite aperture dimensions that are representative of typical experimental conditions.

## 6.8 Conclusion

In this paper, we have shown that sparse regularization based of the use of the $\ell^1$-norm significantly reduces causal errors typically found in the reconstruction of causal signals. For signals that have a higher level of sparsity, this considerably reduces the global RMS error in comparison to standard Tikhonov regularization. In the studied cases, the global RMS error is reduced by half when half of the reconstructed field values are sparse and by a factor three when the sparsity of the reconstructed field reaches 70 %. At time and positions for which the reconstruction is non-null, $\ell^2$ regularization only slightly outperforms the $\ell^1$ method. It was shown that the Pareto frontier curve adequately predicts the optimal value of the regularization parameter with the TSD formulation. The difference in relative RMS error obtained with optimal regularization and with the Pareto frontier curve remains below 1 % for SNR of 30 and 40 dB, below 2 % for SNR of 20 dB and below 6 % for SNR of 10 dB. Additionally, sparse regularization was successfully applied to the reconstruction of experimental transient signal obtained from the impact of a metallic rod on a PMMA plate.

In the past years, many efforts were invested in finding methods to predict the optimal Tikhonov parameter. Since the performance of $\ell^1$ regularization is promising, future work could focus on investigating the application of prediction methods to such regularization method. In addition, performance of the elastic net method which consists in a combination of both $\ell^1$ and $\ell^2$ regularization methods could be investigated. It would profit from the advantage of both sparse and Tikhonov as an improved reconstruction could lie in such combination.

## 6.9 Appendix A

### 6.9.1 Piston model

An analytical model to determine the transient radiation of a circular baffled piston was developed by Stepanishen [45]. The model is based on Green's function formulation and is exact in both the near-field and the far-field, and it is applicable to any displacement of the piston. Since the displacement of the piston is uniform in space, Stepanishen shows that the pressure field with respect to time and position $r$ is expressed as:

$$p(r,t) = \rho_0 h(r,t) * \frac{\partial^2 x(t)}{\partial t^2}. \tag{6.12}$$

The $*$ symbol denotes the time domain convolution of the acceleration with the impulse response $h(r,t)$, and $\rho_0$ is the density of air. When piston radius $a > r$, the impulse response function is defined as:

$$h(r,t) = \begin{cases} 0, & ct < l \\ c, & l < ct < R_{min} \\ \frac{c}{\pi} cos^{-1}\left(\frac{(ct)^2 - l^2 + r^2 - a^2}{2r\sqrt{(ct)^2 - l^2}}\right), & R_{min} < ct < R_{max} \\ 0, & ct > R_{max} \end{cases} \tag{6.13}$$

When $a < r$, the impulse response function becomes:

$$h(r,t) = \begin{cases} 0, & ct < R_{min} \\ \frac{c}{\pi} cos^{-1}\left(\frac{(ct)^2 - l^2 + r^2 - a^2}{2r\sqrt{(ct)^2 - l^2}}\right), & R_{min} < ct < R_{max} \\ 0, & ct > R_{max} \end{cases} \tag{6.14}$$

where $R_{min}$ and $R_{max}$ are respectively the shortest and longest distances from the observation point to the circumference of the piston:

$$\begin{aligned} R_{min} &= \sqrt{z^2 + (a-r)^2}, \\ R_{max} &= \sqrt{z^2 + (a+r)^2}, \end{aligned} \tag{6.15}$$

and $l$ is the projection of $r$ over the $z$ axis. In this analysis, an under-damped sinusoidal piston motion profile is used:

$$x(t) = \mathrm{A} \cdot \sin(2\pi f t) \cdot e^{-\zeta t}, \tag{6.16}$$

$$\frac{\partial^2 x(t)}{\partial t^2} = Ae^{-\zeta t}\left(-\omega^2 sin(\omega t - \pi/2) - 2\zeta\omega cos(\omega t - \pi/2) + \zeta^2 sin(\omega t - \pi/2)\right), \tag{6.17}$$

where $\zeta$ is the decay coefficient. The physical parameters involved in the calculation are presented in Tableau 6.1.

*Tableau 6.1: Physical parameters of the transient piston*

| *Parameters* | *Values* |
|---|---|
| Piston radius | $a = 0.20$ m |
| Sinusoidal frequency | $f = 500$ Hz |
| Decay coefficient | $\zeta = 1000\text{s}^{-1}$ |
| Amplitude | $A = 0.0001$m |
| Speed of sound | $c = 343$ m/s |
| Air density | $\rho_0 = 1.2$ kg/m$^3$ |
| Measuring distance | $z_0 + d = 0.05$ m |

# CHAPITRE 7 DISCUSSION GÉNÉRALE

Ce chapitre présente une discussion générale sur les principales contributions originales développées dans cette thèse. On discute de l'atteindre les objectifs fixés et de la place des travaux par rapport à l'état de l'art en holographie temporelle.

## 7.1 Atteinte des objectifs

L'ensemble des objectifs énoncés au chapitre 3 a été atteint. L'objectif principal de la thèse est de « développer une formulation d'holographie acoustique en champ proche pour des sources non-stationnaires et causales qui permet de déterminer avec précision le champ sonore ». À travers les articles 1 à 3, la formulation temps-espace est effectivement développée. Pour le problème direct, nous démontrons que la méthode permet de déterminer le champ avec une très bonne précision : l'erreur est d'environ 3 % pour une distance de propagation allant jusqu'à 55 cm. L'erreur reste environ constante avec la distance de propagation. La principale source d'erreur en propagation est due à la troncature du champ mesuré et de la fonction de Green. Si les fenêtres spatiales et temporelles sont suffisamment grandes, cette erreur de troncature devient négligeable. Pour le problème inverse, l'objectif est également atteint puisque la formulation temps-espace permet de reconstruire le champ avec une bonne précision malgré la nature mal-posée du problème. Cependant, la formulation reste imparfaite puisque l'application du *zero-padding* au champ mesuré engendre des erreurs dans le cas où le champ est tronqué. Nous avons démontré que l'acquisition dans l'espace tridimensionnel engendre nécessairement cette troncature, et donc une erreur dans la reconstruction. Cette erreur peut cependant être réduite significativement grâce à l'algorithme *patch*, ou bien en considérant une très grande fenêtre d'acquisition. L'application d'une régularisation adéquate est également un aspect important à l'obtention d'une reconstruction précise. Une application des méthodes $\ell^1$ et $\ell^2$ suivant les recommandations issues de notre analyse permet d'appliquer une régularisation qui s'approche de la solution optimale. Nous concluons donc que l'objectif principal est atteint.

Ces travaux permettent également l'atteinte du sous-objectif *i*, puisque la formulation temps-espace est basée sur l'utilisation de la convolution linéaire, et donc, supprime complètement les erreurs de recouvrement. La convolution linéaire est obtenue en ajoutant un *zero-padding* aux signaux convolués et nous avons également proposé deux nouveaux algorithmes qui permettent son calcul

de manière plus efficace en propagation. En rétro-propagation, les erreurs de recouvrement sont également supprimées, puisque l'inverse de la matrice de convolution linéaire est utilisé.

Dans le même ordre d'idée, la formulation temps-espace permet de supprimer les erreurs de fuites spectrales, puisqu'elles sont causées par l'échantillonnage de la fonction de Green dans le domaine de Fourier. Avec la formulation proposée, nous évitons ce problème en échantillonnant directement dans le domaine temps-espace. Le sous-objectif *ii* est donc atteint.

Les performances et les limites des méthodes de régularisation $\ell^2$ et $\ell^1$ ont été analysées en profondeur dans les articles 2 et 3. Ces méthodes permettent de réduire grandement l'erreur de déconvolution, malgré que celle-ci reste considérable si le niveau de bruit ou la distance de rétro-propagation est importante. La régularisation devient également inefficace si le paramètre de régularisation $\lambda$ n'est pas choisi adéquatement. Nous avons démontré que la méthode $\ell^1$ produit une plus petite erreur RMS que la méthode $\ell^2$ si le signal reconstruit contient une grande proportion de valeurs nulles. Cependant, outre ce régime pour lequel la solution est nulle, la méthode $\ell^2$ performe (légèrement) mieux. De plus, la méthode $\ell^1$ requiert un nombre d'opération beaucoup plus grand que la méthode $\ell^2$ : en termes de temps de calcul, il y a environ un facteur cent entre les deux méthodes. Nous considérons donc le sous-objectif *iii* atteint.

Les performances des méthodes de prédiction du paramètre de régularisation ont été évaluées pour les deux méthodes de régularisation, conformément à l'objectif *iv*. Dans le cas de la méthode de régularisation $\ell^2$, trois méthodes ont été testées. Avec le modèle du piston bafflé, la méthode de validation croisée généralisée donne les meilleures performances, suivie de la méthode bayésienne. Cependant, les performances de ces méthodes dépendent du niveau de bruit et de la mesure, il serait donc intéressant de tester ces méthodes avec d'autres types de source : les conclusions pourraient alors différer. Pour la régularisation $\ell^1$ nous avons démontré que la courbe de Pareto permet de prédire adéquatement le paramètre de régularisation. À notre connaissance, il s'agit d'une première application de cette méthode au problème inverse en NAH.

L'ensemble des calculs de convolution et de déconvolution aux articles 1 à 3 se base sur l'utilisation de la transformée de Fourier rapide. Cela permet de conserver une grande efficacité de calcul, ce qui souscrit à l'objectif *v*. De plus, les algorithmes développés aux articles 4 et 5 permettent une amélioration de l'efficacité de calcul de la convolution linéaire, qui est utilisée pour le problème direct ainsi que dans l'algorithme d'optimisation convexe de la méthode $\ell^1$.

Finalement, des validations sont effectuées à partir du modèle de piston pour les articles 1 à 3 et des validations expérimentales sont présentées aux articles 1 et 3. L'objectif *vi* est donc atteint.

## 7.2 Impact des travaux sur l'état de l'art en holographie acoustique temporelle

Un principal impact de ces travaux de thèse sur l'état de l'art est la mise en valeur de la convolution linéaire dans le NAH. En effet, bien que le phénomène de *wrap-around* soit énoncé dans la littérature, sa cause n'est pas établie en profondeur, malgré que ce soit une notion de base en traitement du signal. En effet, le terme « convolution linéaire » n'apparait pas dans la littérature, bien que celle-ci soit appliquée dans certains cas [6, 7]. De plus, les solutions proposées pour remédier au *wrap-around*, qui sont décrites dans certains des plus importants ouvrages [2, 3], ne sont pas efficaces. Une compréhension plus approfondie du phénomène est mise de l'avant par nos travaux. Cela s'explique en partie par le fait que le *wrap-around* est mis en évidence plus explicitement lorsque le signal est représenté dans le domaine temporel. En effet, dans le domaine temps-espace, l'interprétation du champ sonore peut être plus évidente, puisqu'il est possible d'observer la propagation des fronts d'ondes. L'erreur de *wrap-around*, qui survient souvent près des frontières du domaine spatial fini, devient alors éminente. Cela est d'autant plus manifeste lorsque les signaux sont causaux.

Les bienfaits de la convolution linéaire sont bien mis de l'avant pour le problème direct dans l'article 1. En effet, il est clair qu'avec ce type de convolution, l'erreur obtenue est réduite de manière significative. C'est également le cas pour le problème inverse tel que décrit dans l'article 2. De plus, la formulation matricielle de la convolution linéaire permet de mieux expliquer le fonctionnement de l'algorithme *patch* : on fait un lien entre le champ est extrapolé sur le plan de mesure et le fait que les signaux convolués linéairement comportent deux fois plus de valeurs pour chaque dimension. Également, la formulation inverse basée sur la convolution linéaire pourrait stimuler la recherche sur l'inversion de très grande matrice bloc-Toeplitz parcimonieuses, ce qui permettrait ultimement de s'affranchir de l'algorithme patch et de certains problèmes de troncature. Rappelons que ces problèmes ne sont pas propres au NAH temporel mais également au cas stationnaire, puisque la convolution linéaire devrait être appliquée pour les domaines spatiaux. Bref, les articles 1 et 2 contribueront certainement à intéresser la communauté scientifique à

considérer l'application de cette convolution dans leurs propres travaux en NAH. Il est également intéressant de constater que la perte d'efficacité de calcul de la convolution linéaire par rapport à la convolution circulaire en NAH est amoindrie en utilisant les algorithmes présentés aux articles 5 et 6.

Par rapport à l'état de l'art, une contribution significative de l'article 2 est la comparaison des méthodes de prédiction du paramètre de régularisation. En effet, la méthode bayésienne récemment apparue dans la littérature en NAH est comparée avec deux méthodes plus standards : le GCV et la courbe en L. Contrairement aux résultats de Pereira [30] pour l'holographie stationnaire, on observe étonnamment que le GCV surpasse la méthode bayésienne avec le NAH temporel, cela malgré que le GCV est reconnu pour donner des résultats plutôt mitigés lorsqu'appliqué au NAH [56]. La raison de cette divergence n'est pas claire : elle pourrait être due à la nature des sources étudiées ou bien à l'implémentation du GCV. Nous espérons que nos résultats permettront d'instiguer des analyses plus approfondies sur le sujet. Nous sommes convaincus que l'approche bayésienne permet des améliorations importantes en NAH, notamment en termes d'interprétation du problème inverse, et il serait intéressant de mieux définir les limites de cette méthode.

Finalement, une autre contribution importante à l'état de l'art est l'application de la régularisation parcimonieuse au problème inverse. Les applications de cette méthode au NAH sont récentes et portent seulement sur la formulation stationnaire. L'interprétation de la parcimonie pour des signaux représentés dans le domaine des fréquences $\omega$ est différente de celle des signaux temporels. L'originalité de notre approche est d'avoir considéré cette méthode pour des sources causales pour lesquelles la parcimonie est assurée. Nous croyons que ces idées pourraient avoir un impact durable sur la recherche en NAH temporel.

# CHAPITRE 8 CONCLUSION ET RECOMMANDATIONS

Ce chapitre résume les principales contributions de cette thèse, expose les limites et contraintes des concepts scientifiques développés et présente certaines recommandations ainsi que des idées de travaux futurs.

## 8.1 Contributions

### 8.1.1 Contributions scientifiques originales

La principale contribution de ce projet de thèse est la méthode d'holographie temps-espace qui permet de déterminer le champ sonore sans erreurs de recouvrement ou de fuites spectrales. Grâce à cette méthode, il est possible de résoudre le problème direct, un problème bien posé mathématiquement, sans que l'erreur augmente de manière significative avec l'augmentation de la distance de propagation. Ainsi, au lieu de proposer une manière d'atténuer ces erreurs, la méthode proposée résout de manière fondamentale et définitive les principaux problèmes des formulations d'holographie temporelle. La principale erreur reste les inévitables troncatures de la fonction de Green et du champ mesuré, cependant, ces deux paramètres sont a priori contrôlables par l'utilisateur.

Pour le problème inverse, les performances des principaux algorithmes de régularisation ont été testées avec la méthode proposée. La régularisation de Tikhonov donne les meilleurs résultats lorsque le signal n'est pas parcimonieux et la méthode de validation croisée généralisée permet de prédire avec la plus grande précision le paramètre de régularisation. Lorsque le champ est parcimonieux, comme c'est le cas pour des sources causales, la méthode de régularisation parcimonieuse permet une meilleure reconstruction du champ là où son amplitude est nulle. La courbe de Pareto permet une prédiction adéquate du paramètre de régularisation pour la méthode parcimonieuse. Ainsi, grâce à l'analyse proposée, le formalisme temps-espace est applicable à la résolution du problème inverse. L'erreur de régularisation reste tout de même significative dans bien des configurations, et ce, malgré une application optimale des algorithmes de régularisation. L'utilisateur doit rester vigilant dans le choix de ces paramètres expérimentaux, notamment en ce qui concerne la taille de l'antenne et la distance de rétro-propagation. Les conséquences de la taille finie de l'antenne peuvent tout de même être amoindries en utilisant l'algorithme *patch*.

Un algorithme pour obtenir la convolution linéaire à partir de la NLT a été établi. Cet algorithme repose sur un coefficient d'atténuation et une formule a été proposée pour déterminer la valeur de ce coefficient de manière à minimiser l'erreur. Cette méthode permet d'obtenir chacune des moitiés du résultat de la convolution linéaire de manière séparée, ce qui permet une économie de calcul pour certaines applications.

Finalement, nous proposons un algorithme de calcul de la convolution linéaire multidimensionnelle basé sur la réutilisation de certains coefficients de régularisation de la GDFT. Cette méthode permet une réduction additionnelle du nombre d'opérations d'environ 20 % par rapport à la méthode de GDFT standard et d'environ 50 % par rapport à la méthode du *zero-padding*. Son application dans le contexte de la méthode d'holographie acoustique temporelle est avantageuse, puisqu'elle permet de réduire les temps de calcul significativement. Cette méthode est également compatible avec l'algorithme d'optimisation convexe utilisé pour la régularisation parcimonieuse.

### 8.1.2 Publications

Ces travaux doctoraux mènent à la soumission de cinq articles à des journaux de premier plan en acoustique et en traitement du signal, dont deux articles ont déjà été publiés. Trois articles constituent le corps de la présente thèse, et les deux autres sont présentés en annexe. Les travaux ont également mené à la publication et à la présentation de quatre articles de conférence à des conférences internationales (ICSV19 [72], ICVSV20 [73], ITCAM 2016 [74] et ICA 2016 [75]), ainsi qu'à la production de cette thèse.

### 8.1.3 Encadrement

Les travaux doctoraux ont mené à l'encadrement de trois stagiaires de premier cycle par l'auteur de cette thèse. Les stagiaires (Dmitri Fedorov, Jeremy Pinto, Marc Provost) ont participé au développement de la partie expérimentale du projet et à la prise de mesures. Ceci leur a permis de les initier à la recherche et de contribuer à leur formation d'ingénieur.

### 8.1.4 Rayonnement du laboratoire

Le projet doctoral a permis de faire rayonner le laboratoire au-delà de ses portes. Outre la participation aux conférences, les principales contributions en ce sens sont :

1. L'obtention du prix « la preuve par l'image » de l'ACFAS (2013),

2. La vulgarisation des travaux de recherche à l'émission « Le code Chastenay » (2014), ainsi que pour une vidéo promotionnelle pour Polytechnique Montréal,
3. L'établissement d'une première collaboration étudiante avec le Laboratoire d'acoustique de l'Université du Maine (LAUM) sous la forme d'un stage de recherche de six mois (2015).

### 8.1.5 Legs

Finalement, deux héritages du projet sont le montage expérimental qui pourra contribuer à la suite des travaux, ainsi que les multiples codes MATLAB utilisés pour les calculs de convolution et de régularisation.

## 8.2 Limites et contraintes

Bien que la méthode développée ait plusieurs avantages importants, celle-ci comporte également certaines limitations.

Premièrement, le formalisme a été développé pour une géométrie planaire : il pourrait être intéressant de développer des formulations équivalentes pour les géométries sphériques et cylindriques. Ces géométries présentent une symétrie circulaire selon certaines dimensions, l'effet de *wrap-around* est alors désirable dans ces cas et la convolution circulaire devrait être appliquée. L'opérateur prendrait alors la forme d'une convolution tridimensionnelle où la convolution linéaire serait appliquée pour certaines dimensions et la convolution circulaire le serait pour les autres. Un échantillonnage de la fonction de Green dans le domaine temps-espace serait tout de même requis pour éviter les problèmes de fuites spectrales.

Une autre limitation de la formulation proposée est que la discrétisation doit être uniforme selon chacun des trois domaines $(x, y, t)$. De plus, il est nécessaire que cette même discrétisation soit conservée dans la reconstruction. La formulation proposée possède cette caractéristique, puisqu'elle tire profit de l'utilisation de la transformée de Fourier rapide 3-D qui réduit de manière très significative les coûts de calcul. Il serait possible de reformuler le problème de manière à outrepasser cette contrainte, cependant, la matrice correspondante comporterait un très grand nombre de valeurs, et son inversion nécessiterait certainement des ressources de calcul très

importantes. Ce problème serait simplifié si la fonction de Green était séparable, c'est-à-dire qu'elle peut s'exprimer comme $g(x,y,t) = g_1(x) \cdot g_2(y) \cdot g_3(t)$, ce qui n'est pas le cas.

Dans un même ordre d'idée, la circularisation de la matrice $G$, qui est nécessaire au calcul de son inverse avec la transformée de Fourier rapide, génère des erreurs de troncature pour le problème inverse, tel que discuté dans l'article 2. Celles-ci peuvent être réduites avec l'algorithme *patch*, mais l'idéal serait de pouvoir inverser la matrice Toeplitz correspondante au champ mesuré, et ainsi s'affranchir de ces erreurs. Il existe certains algorithmes permettant une inversion plus rapide de ces matrices, mais leur applicabilité au problème de NAH n'a pas été testée. De plus, malgré l'utilisation de la transformée de Fourier rapide, certains calculs restent assez demandant, notamment dans le cas de la résolution du problème convexe pour la régularisation $\ell^1$. L'efficacité pourrait être améliorée en optimisant l'algorithme de résolution ou en programmant celui-ci dans un langage de programmation plus performant.

Une autre limite de notre analyse porte sur le fait que la performance des algorithmes de régularisation n'a été déterminée que pour un seul type de source simulée : le piston bafflé. Cette source comporte plusieurs avantages qui justifient son utilisation : une connaissance de la solution analytique de son champ, la facilité de son implémentation, son applicabilité a de multiples profils de déplacement, sa ressemblance avec la plaque impactée analysée expérimentalement, etc. Cependant, les performances des algorithmes de régularisation dépendent du champ mesuré et il serait intéressant de faire une analyse plus approfondie en considérant d'autres types de sources. Par exemple, on pourrait considérer une source qui souscrit mieux aux hypothèses bayésiennes de distribution gaussienne, il serait alors plausible de croire que la méthode de prédiction du coefficient de régularisation basée sur cette hypothèse aurait de meilleures performances par rapport à la validation croisée généralisée.

Finalement, la méthode proposée s'applique seulement en champ libre et dans un milieu non-dispersif. Il serait intéressant de développer une adaptation pour pallier ces limitations, notamment dans le contexte de l'acoustique sous-marine.

## 8.3 Recommandations et travaux futurs

Tel que discuté ci-dessus, la méthode proposée, bien qu'elle permette d'améliorer significativement la détermination du champ sonore, présente tout de même certaines limitations.

Des travaux futurs pourraient viser à pallier ces limitations. Néanmoins, nous dirigeons l'attention du lecteur sur certaines recommandations de travaux futurs qui, nous croyons, ont un intérêt particulier.

Puisqu'il est maintenant possible de reconstruire le champ acoustique non-stationnaire à la source avec précision, il serait intéressant d'utiliser ces reconstructions pour caractériser des structures mécaniques. Un intérêt d'appliquer cette analyse à une excitation non-stationnaire est qu'il est possible de traiter le problème sur tout le spectre en simultané. On peut ainsi, impacter une plaque homogène et déterminer son champ à la source : ceci met en évidence certaines caractéristiques, comme par exemple, la vitesse de propagation des ondes dans la plaque selon la fréquence, ce qui permet de caractériser le matériau de la plaque. Il est aussi possible de localiser un défaut dans la plaque grâce au champ sonore. Une étude préliminaire a été effectuée en ce sens et les résultats ont été présentés au congrès ICA en 2016 [75]. À la Figure 8.1, on présente l'évolution spatiale du champ sonore obtenu expérimentalement reconstruit au plan de la source pour trois temps donnés.

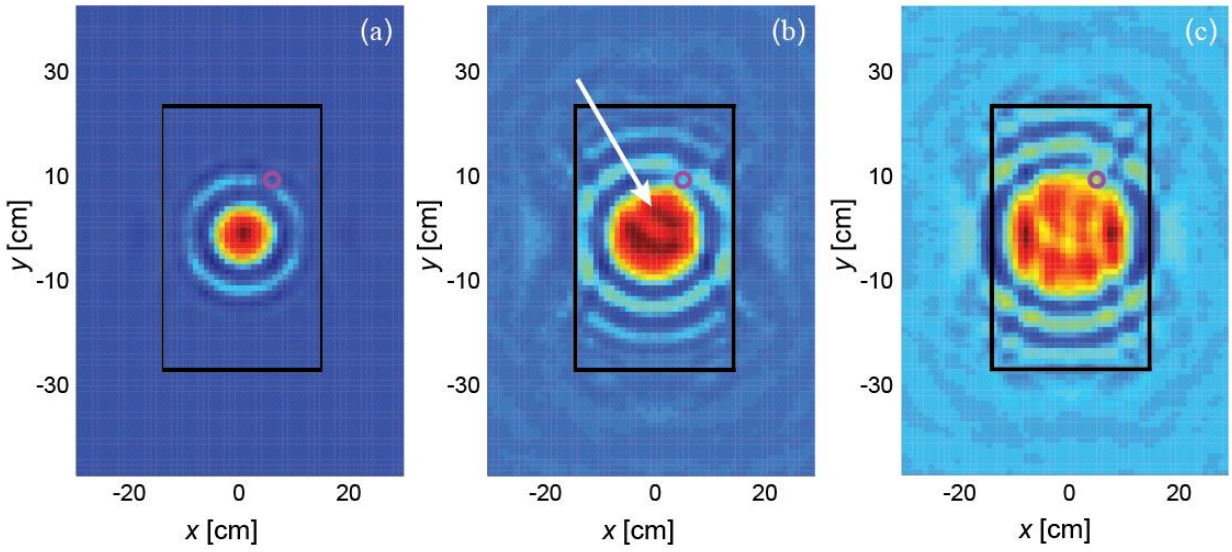


*Figure 8.1 : Représentation spatiale du champ transitoire reconstruit au plan de la source. (a)* $t = 0.17$ *ms, (b)* $t = 0.25$ *ms et (c)* $t = 0.34$ *ms. La flèche en (b) indique la présence de fronts d'onde réfléchis. Tiré de [75]*

Le champ est issu d'une plaque de plexiglass impactée à son centre et dont les dimensions sont délimitées par le carré noire. Celle-ci comporte un trou de 3.5 cm de diamètre dont la position est indiquée par le cercle mauve. En étudiant le champ reconstruit, on remarque que le front d'onde est réfléchi à cause du changement d'impédance à la position du défaut (notamment désigné par

une flèche à la Figure 8.1 (b)). On peut ainsi imaginer que l'application d'un algorithme de traitement d'image ou d'intelligence artificielle permettrait d'automatiser la caractérisation de tels défauts.

Jusqu'à maintenant, ce genre d'application à l'holographie acoustique n'a pas été très exploité dans la littérature. La méthode développée dans cette thèse permettrait d'effectuer une telle analyse dans le domaine temps-espace avec une plus grande précision et favoriserait sans doute l'obtention de résultats intéressants.

Il serait également intéressant de poursuivre les travaux sur la régularisation du problème inverse, notamment en analysant les performances de l'*elastic-net*, qui est une méthode hybride entre la régularisation $\ell^1$ et $\ell^2$. Cette méthode permettrait de tirer avantage des propriétés de la régularisation $\ell^1$ dans le régime pour lequel la solution est nulle, et de celles de la régularisation $\ell^2$ lorsque l'amplitude de la solution n'est pas nulle. Il serait aussi intéressant de développer la méthode de prédiction du paramètre de régularisation avec l'approche bayésienne expérimentale de Pereira et Antoni [30] pour la régularisation $\ell^1$. Cela permettrait d'avoir une alternative à la méthode de la courbe de Pareto dans les cas où celle-ci ne donne pas de bons résultats.

Bref, ces travaux de thèse ont permis de répondre à plusieurs questions épineuses et leur contribution à l'état de l'art en holographie acoustique temporelle est significative. Ces réponses ont permis de mettre en lumière d'autres aspects du domaine de recherche qui méritent d'être explorés plus en profondeur. Nous croyons que les conclusions de cette thèse permettront ainsi de définir certaines avenues de recherche pour les années à venir.

# BIBLIOGRAPHIE

[1] E. G. Williams et J. D. Maynard, «Holographic imaging without the wavelength resolution limit,» *Physical Review Letters*, vol. 45, n° %17, pp. 554-557, 1980.

[2] E. G. Williams, Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography, New York: Academic Press, 1999, pp. 1-306.

[3] J. D. Maynard, E. G. Williams et Y. Lee, «Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH,» *Journal of the Acoustical Society of America*, vol. 78, pp. 1395-1413, 1985.

[4] P. K. Morse et K. U. Ingard, Theoretical Acoustics, McGraw-Hill, 1968.

[5] W. L. Briggs et V. E. Henson, The DFT: An owner's manual for the discrete Fourier transform, Philadelphia: Society for Industrial and Applied Mathematics, 1995, pp. 1-434.

[6] V. Grulier, S. Paillasseur, J.-H. Thomas et J.-C. Pascal, «Forward propagation of time evolving acoustic pressure : Formulation and investigation of the impulse response in time-wavenumber domain,» *Journal of the Acoustical Society of America*, vol. 126, n° %15, pp. 2367-2378, 2009.

[7] O. De la Rochefoucauld, M. Melon et A. Garcia, «Time domain holography: Forward projection of similated and measured sound pression fields,» *Journal of the Acoustical Society of America*, vol. 116, n° %11, pp. 142-153, 2004.

[8] J.-F. Blais et A. Ross, «Backward propagation of sound fields radiated by impacted plates using a transient acoustical holography approach,» chez *38th International Congress and Exposition on Noise Control Engineering 2009*, Ottawa, 2009.

[9] P. C. Hansen, Rank-Deficient and Discrete Ill-Posed Problems, Philadelphia: SIAM, 1998.

[10] P. C. Hansen, «Deconvolution and regularization with Toeplitz matrices,» *Numerical algorithms,* vol. 29, pp. 323-378, 2002.

[11] C. W. Elmore et M. A. Heald, Physics of Waves, New York: Dover, 1969.

[12] T. B. Hansen et A. D. Yaghjian, «Formulation of time-domain planar near-field measurement without probe correction,» Rome Laboratory In-House Report RL-TR-93-210, 1993.

[13] A. D. Pierce, Acoustics : an introduction to its physical principles and applications, New York: McGraw-Hill, 1981, pp. 8-11.

[14] Arfken, Weber et Harris, Mathematical Methods for Physicists, 7th edition, Waltham, MA: Elsevier, 2013.

[15] S. A. Hambric, «Structural acoustics tutorial - part 1: Vibrations in structures,» *Acoustics today,* pp. 21-33, 2006.

[16] S. A. Hambric et J. B. Fahnline, «Structural acoucstics tutorial - part 2: Sound-structure interactions,» *Acoustics Today,* pp. 9-27, 2007.

[17] H. J. Nussbaumer, Fast Fourier Transform and Convolution Algorithms, Berlin: Springer-Verlag, 1982.

[18] B. Hunt, «A matrix theory proof of the discrete convolution theorem,» *IEEE Transactions on Audio and Electroacoustics,* vol. 19, n° %14, pp. 285-288, 1971.

[19] P. J. Davis, Circulant matrices, New York: Chelsea Publishing, 1994.

[20] J. Martinez, R. Heusdens et R. Hendriks, «A generalized poisson summation formula and its application to fast linear convolution,» *IEEE Signal Processing letters,* vol. 18, pp. 501-504, 2011.

[21] M. Narasimha, «Linear convolution using skew-cyclic convolutions,» *IEEE Signal Processing letters*, vol. 14, pp. 173-176, 2007.

[22] A. V. Oppenheim et R. W. Schafer, Digital Signal Processing, Prentice-Hall, 1975.

[23] R. Tibshirani, «Regression shrinkage and selection via the lasso,» *J. Royal. Statist. Soc B.,* vol. 58, n° %11, pp. 267-288, 1996.

[24] S. Boyd et L. Vandenberghe, Convex Optimization, New York, NY, USA: Cambridge University Press , 2004.

[25] P. C. Hansen et D. P. O'Leary, «The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems,» *SIAM J. Sci. Comput.,* vol. 14, n° %16, p. 1487–1503, 1993.

[26] E. Van Den Berg et M. P. Friedlander, «Probing the Pareto frontier for basis pursuit solutions,» *SIAM J. Sci. Comput.,* vol. 31, n° %112, pp. 890-912, 2008.

[27] J. Nasehi Tehrani, A. McEwan, C. Jin et A. van Schaik, «L1 regularization method in electrical impedance tomography by using the L1-curve (Pareto frontier curve),» *Applied Mathematical Modelling,* vol. 36, pp. 1095-1105, 2012.

[28] P. Craven et G. Wahba, «Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation,» *Numerical mathematics,* vol. 31, pp. 377-403, 1979.

[29] G. H. Golub, M. Heath et G. Wahba, «Generalized cross-validation as a method for choosing a good ridge parameter,» *Technometrics,* vol. 21, n° %12, pp. 215-223, 1979.

[30] A. Pereira, J. Antoni et Q. Leclère, «Empirical Bayesian regularization of the inverse acoustic problem,» *Applied Acoustics,* vol. 97, pp. 11-29, 2015.

[31] J. Antoni, «A Bayesian approach to sound source reconstruction: Optimal basis, regularization and focusing,» *Journal of the Acoustical Society of America*, vol. 131, n° %14, p. 2873–2890, 2012.

[32] J. Hald, «Time domain acoustical holography and its applications,» *Sound and Vibrations*, vol. 35, n° %12, pp. 16-25, 2001.

[33] F. Deblauwe, J. Leiridan, J. L. Chauray et B. Beguet, «Acoustic holography in transient conditions,» chez *International congress on sound and vibrations*, Copenhagen, 1999.

[34] J.-H. Thomas, V. Grulier, S. Paillasseur et J.-C. Pascal, «Real-time near-field acoustic holography for continuously visualizing nonstationary acoustic fields,» *Journal of the Acoustical Society of America*, vol. 128, n° %16, pp. 3554-3567, 2010.

[35] M. Forbes, S. Letcher et P. Stepanishen, «A wave vector, time-domain method of forward projecting time-dependent pressure fields,» *Journal of the Acoustical Society of America*, vol. 90, n° %15, pp. 2782-2793, 1991.

[36] S. Paillasseur, J.-H. Thomas et J.-C. Pascal, «Regularization for improving the deconvolution in real-time near-field acoustic holography,» *Journal of the Acoustical Society of America*, vol. 129, n° %16, pp. 3777-3787, 2011.

[37] T. Le Magueresse, J.-H. Thomas, J. Antoni et S. Paillasseur, «Instantaneous Bayesian regularization applied to real-time near-field acoustic holography,» *Journal of the Acoustical Society of America*, vol. 142, n° %12, pp. 924-934, 2017.

[38] X.-Z. Zhang, J.-H. Thomas, C.-X. Bi et J.-C. Pascal, «Reconstruction of nonstationary sound fields based on the time domain plane wave superposition method,» *Journal of the Acoustical Society of America*, vol. 132, n° %14, p. 2427–2436, 2012.

[39] S. F. Wu, H. Lu et M. S. Bajwa, «Reconstruction of transient acoustic radiation from a sphere,» *Journal of the Acoustical Society of America*, vol. 117, pp. 2065-2077, 2005.

[40] H. Inoue, M. Kamibaya, K. Kishimoto, T. Shibuya et T. Koizumi, «Numerical Laplace transform and inversion using fast Fourier transform,» *JSME International Journal,* vol. 35, n° %13, pp. 319-324, 1992.

[41] X.-Z. Zhang, C.-X. Bi, Y.-B. Zhang et L. Xu, «Transient nearfield acoustic holography based on an interpolated time-domain equivalent source method,» *Journal of the Acoustical Society of America,* vol. 130, n° %13, pp. 1430-1440, 2011.

[42] C.-X. Bi, L. Geng et Z.-X. Zhang, «Cubic Spline interpolation-based time-domain equivalent source method for modeling transient acoustic radiation,» *Journal of Sound and Vibration,* vol. 332, p. 5939–5952, 2013.

[43] G. Chardon, L. Daudet, A. Peillot, F. Ollivier, N. Bertin et R. Gribonval, «Nearfield Acoustic Holography using sparsity and compressive sampling principles,» *Journal of the Acoustical Society of America,* vol. 132, n° %13, pp. 1521-1534, 2012.

[44] E. Fernandez-Grande, A. Xenaki et P. Gersoft, «A sparse equivalent source method for near-field acoustic holography,» *Journal of the Acoustical Society of America,* vol. 141, n° %11, pp. 532-542, 2016.

[45] P. R. Stepanishen, «Transient radiation from pistons in an infinite baffle,» *Journal of the Acoustical Society of America,* vol. 49, n° %15, pp. 1629-1638, 1970.

[46] J.-F. Blais et A. Ross, «Forward projection of transient sound pressure fields radiated by impacted plates using numerical Laplace transform,» *Journal of the Acoustical Society of America,* vol. 125, n° %15, pp. 3120-3128, 2009.

[47] J.-M. Attendu et A. Ross, «Method for finding optimal exponential decay coefficient in numerical Laplace transform for application to linear convolution,» *Signal Processing,* vol. 130, pp. 47-56, 2017.

[48] S. Pan, W. Jiang, S. Xiang et X. Lui, «An interpolated time-domain equivalent source method for modeling transient acoustic radiation over a mass-like plane based on the transient half-space Green's function,» *Wave Motion,* vol. 51, pp. 1273-1287, 2014.

[49] M. Fedia, J. Antoni, F. Taher, V. Lafranchi et H. Mohamed, «Wavelet-based non-stationary near-field acoustical holography,» *Applied Acoustics,* vol. 74, pp. 1226-1233, 2013.

[50] C. Radhakrishnan et W. K. Jenkins, «The 2-D modulated discrete Fourier transform for 2-D fast convolution and digital filtering,» chez *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2011.

[51] R. Scholte, I. Lopez, B. N. Roozen et H. Nijmeijer, «Truncated aperture extrapolation for Fourier-based near-field acoustic holography by means of border-padding,» *Journal of the Acoustical Society of America,* vol. 125, n° %16, pp. 3844-3854, 2009.

[52] L. M. Milne-Thomson, The calculus of finite differences, 2nd éd., New York: Chelsea, 1971, pp. 154-196.

[53] J. J. Stickel, «Data smoothing and numerical differentiation by a regularization method,» *Computers and Chemical Engineering,* vol. 34, pp. 467-475, 2010.

[54] J.-M. Attendu et A. Ross, «Nearfield acoustical holography without wraparound error and spectral leakage for nonstationary forward propagation,» *Journal of the Acoustical Society of America,* vol. 1, n° %11, pp. 1000-1021, 2017.

[55] L. Geng, X.-Z. Zhang et C.-X. Bi, «Reconstruction of transient vibration and sound radiation of an impacted plate using time domain plane wave superposition method,» *Journal of Sound and Vibrations,* vol. 344, p. 114–125, 2015.

[56] E. G. Williams, «Regularization methods for near-field acoustical holography,» *Journal of the Acoustical Society of America,* vol. 110, n° %14, pp. 1976-1988, 2001.

[57] M. Wax et T. Kailath, «Efficient Inversion of Toeplitz-Block Toeplitz,» *IEEE transactions on acoustics, speech, and signal processing,* vol. 31, n° %15, pp. 1218-1221, 1983.

[58] A. E. Yagle, «A fast algorithm for Toeplitz-block-Toeplitz linear system,» chez *IEEE International Conference on Acoustics, Speech, and Signal Processing,* Salt Lake City, 2001.

[59] G. H. Golub et C. F. Van Loan, Matrix computations, Baltimore: Johns Hopkins University Press, 1983, p. 476 p..

[60] J. Hald, «Reduction of Spatial Windowing effects in acoustical holography,» chez *inter-noise 94,* Yokohama, 1994.

[61] E. G. Williams, «Continuation of Acoutic near-fields,» *Journal of the Acoustical Society of Americal,* vol. 113, n° %13, pp. 1273-1281, 2002.

[62] E. G. Williams et B. H. Houston, «Fast Fourier transform and sigular value decomposition formulations for patch nearfield acoustical holography,» *Journal of the Acoustical Society of America,* vol. 114, n° %13, pp. 1322-1333, 2003.

[63] K. Saijyou et S. Yoshikawa, «Reduction methods of the reconstruction error for large-scale implementation of near-field acoustical holography,» *Journal of the Acoustical Society of America,* vol. 110, n° %14, pp. 2007-2023, 2001.

[64] J. Hald, « Basic theory and properties of statistically optimized near-field acoustical holography,» *Journal of the Acoustical Society of America,* vol. 125, n° %14, pp. 2105-2120, 2009.

[65] P. C. Hansen, «Regularization Tools version 4.0 for Matlab 7.3,» *Numerical Algorithms,* vol. 46, pp. 189-194, 2007.

[66] J.-M. Attendu et A. Ross, «Time domain nearfield acoustical holography with three-dimensional linear deconvolution,» *Journal of the Acoustical Society of America,* To be published.

[67] P. Simard et J. Antoni, «Acoustic source identification: Experimenting the l1 minimization approach,» *Applied Acoustics*, vol. 74, pp. 974-986, 2013.

[68] H. F. Alqadah et N. Valdivia, «Near-field electromagnetic holography for arbitrary surfaces using sparse regularization,» chez *International Conference on Electromagnetics in Advanced Applications (ICEAA)*, Torino, Italy, 2013.

[69] Y. Zhang, J. Yang et W. Yin, «YALL1: Your ALgorithms for L1, online at yall1.blogs.rice.edu,» 2011. [En ligne].

[70] J. Huang, C. R. Berger, S. Zhou et J. Huang, «Comparison of basis pursuit algorithms for sparse channel estimation in underwater acoustic OFDM,» chez *OCEANS IEEE*, Sydney, Australia, 2010.

[71] A. Ross et G. Ostiguy, «Propagation of the initial transient noise from an impacted plate,» *Journal of Sound and Vibrations*, vol. 301, n° %11-2, pp. 28-42, 2007.

[72] J.-M. Attendu et A. Ross, «Investigation of transient near-field acoustical holography using temporal and spatial Laplace transforms,» chez *ICSV19*, Vilnius, 2012.

[73] J.-M. Attendu et A. Ross, «Skew-circulant matrix formulation for transient near-field acoustical holography,» chez *ICSV20*, Bangkok, 2013.

[74] J.-M. Attendu et A. Ross, «Time-domaine near-field acoustical holography as a means to prevent exposition to harmful industrial impact noises,» chez *ICTAM*, Montreal, 2016.

[75] J.-M. Attendu, A. Ross et J.-H. Thomas, «Localization of a defect in an impacted plate using nearfield acoustical holography and time evolution of spatial features,» chez *Internation congress on acoustics*, Buenos Aires, 2016.

[76] E. Butkov, Mathematical physics, Reading, Mass.: Addison-Wesley, 1968, p. 735.

[77] J. K. Goodman, Introduction to Fourier's Optics, Third Edition, Roberts & Compagny, 2005.

[78] D. J. Wilcox, «Numerical Laplace transformation and inversion,» *Int. J. Elect. Eng. Educ.*, vol. 15, pp. 247-265, 1978.

[79] P. Gómez et F. Uribe, «The numerical Laplace transform: An accurate technique for analyzing electromagnetic transients on power system devices,» *International Journal of Electrical Power & Energy Systems*, vol. 31, n° %12-3, pp. 116-123, 2009.

[80] L. M. Wedepohl, «Power system transients: Errors incurred in the numerical inversion of the Laplace transform,» chez *Proc. 26th Midwest Symp. Circuits Systems*, 1983.

[81] A. Ametani et K. Imanishi, «Development of exponential Fourier transform and its application to electrical transients,» chez *Proc. Inst. Elect. Eng.*, 1978.

[82] F. Uribe, J. Naredo et P. G. L. Moreno, «Electromagnetic transients in underground transmission systems through the numerical Laplace transform,» *International Journal of Electrical Power and Energy systems*, vol. 24, n° %13, pp. 215-221, 2002.

[83] Z. Babic et D. Mandic, «A fast algorithm for linear convolution of discrete time signals,» chez *TELSIKS*, Nis, Yugoslavia, 2001.

[84] J. Martinez, R. Heusdens et R. Hendriks, «A generalized Fourier domain: Signal processing framework and applications,» *Signal Processing*, vol. 93, n° %15, pp. 1259-1267, 2013.

[85] P. Moreno et A. Ramirez, «Implementation of the numerical Laplace transform: a review,» *IEEE Transaction on Power Delivery*, vol. 23, n° %14, pp. 2599-2609, 2008.

[86] Titchmarsh, Introduction to the Theory of Fourier Integrals, London: Oxford University Press, 1948.

[87] J. Hilgevoord, Dispersion Relations and Causal Description: An introduction to Dispersion Relations in Field Theory, Amsterdam: North-Holland Publishing Compagny, 1960.

[88] F. Franchetti, M. Puschel, Y. C. S. Voronenko et J. M. F. Moura, «Discrete Fourier transform on multicore,» *IEEE Signal Processing Magasine*, vol. 26, n° %16, pp. 90-102, 2009.

[89] M. Frigo et S. G. Johnson, «FFTW: an adaptative software architecture for the FFT,» *Proceedings of the International Conference on Acoustics, Speech, Signal Processing*, vol. 3, pp. 1381-1384, 1998.

[90] S. J. Day, N. Mullineux et J. R. Reed, «Developments in obtaining transient response using Fourier transforms. Part I: Gibb phenomena and Fourier integrals,» *Int. J. Elect. Eng. Educ.*, vol. 3, pp. 501-506, 1965.

[91] S. J. Day, N. Mullineux et J. R. Reed, «Developments in obtaining transient response using Fourier transforms. Part II: Use of the modified Fourier transform,» *Int. J. Elect. Eng. Educ.*, vol. 4, pp. 31-40, 1966.

[92] D. Gottlieb et C. Shu, «On the Gibbs phenomenon and its resolution,» *SIAM rev.*, vol. 4, n° %1644-668, p. 39, 1994.

[93] D. J. Inman, Engineering Vibrations, Prentice-Hall, 1994.

[94] C. Radhakrishnan et W. K. Jenkins, «Quadratic modified Fermat transform for fast convolution and adaptative filtering,» chez *IEEE International Conference of Acoustics, Speech and Signal Processing*, 2010.

[95] W. Li et A. M. Peterson, «FIR Filtering by the Modified Fermat Number Transform,» *IEEE Transaction on Acoustics, Speech and Signal Processing*, vol. 38, n° %19, pp. 1641-1645, 1990.

[96] C. Radhakirshnan et W. K. Jenkins, «Performance Analysis for 2-D Convolution Implemented with the 2-D Modified Discrete Fourier Transform,» chez *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, 2012.

[97] S. van der Walt, S. C. Colbert et G. Varoquaux, «The NumPy Array: A Structure for Efficient Numerical Computation,» *Computing in Science & Engineering*, vol. 13, pp. 22-30, 2011.

# ANNEXE A : NOTIONS PRÉLIMINAIRES

## A.1 Transformée de Fourier analytique

Soit la fonction analytique $f(t)$, on définit la transformée de Fourier (*TF*) et la transformée de Fourier inverse selon les équations suivantes [76] :

$$\mathcal{F}\{f(t)\} = f(\omega) = \int_{-\infty}^{\infty} f(t) \cdot e^{-i\omega t}\, dt, \tag{A.1}$$

et

$$\mathcal{F}^{-1}\{f(\omega)\} = f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\omega) \cdot e^{i\omega t}\, d\omega, \tag{A.2}$$

où $t$ est la variable temporelle, $\omega = 2\pi f$ est la fréquence angulaire et $i = \sqrt{-1}$. On conserve la même forme pour la transformée spatiale, en substituant la variable temporelle $t$ par une variable spatiale $x$, $y$ ou $z$. La variable fréquentielle correspondante$k_x$, $k_y$ ou $k_z$, est alors la projection du vecteur d'onde $k = \omega/c$ sur le domaine spatial associé, où $c$ est la vitesse de phase.

De ce fait, la *TF* 3-D utilisée en NAH et son inverse s'expriment ainsi :

$$f(k_x, k_y, \omega) = \int_{-\infty}^{\infty} f(x, y, t) \cdot e^{-i(\omega t + k_x x + k_y y)}\, dx dy dt \tag{A.3}$$

et

$$f(x, y, t) = \int_{-\infty}^{\infty} f(k_x, k_y, \omega) \cdot e^{i(\omega t + k_x x + k_y y)}\, dk_x dk_y d\omega. \tag{A.4}$$

On définit par la suite quelques propriétés importantes de la *TF*, dont les preuves relativement simples sont disponibles dans la littérature [76, 5, 14].

### Propriété de convolution

La convolution $h(t)$ de deux fonctions $f(t)$ et $g(t)$ est une opération linéaire et commutative définie comme [76] :

$$h(t) = f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau) \cdot g(t-\tau)\, d\tau. \tag{A.5}$$

Dans le domaine de Fourier, le résultat de la convolution $h(\omega)$ s'exprime comme le produit de la transformée de Fourier des fonctions convoluées [76] :

$$\mathcal{F}\{h(t)\} = h(\omega) = f(\omega) \cdot g(\omega). \tag{A.6}$$

La propriété de convolution établit donc l'équivalence entre l'opération de convolution dans le domaine direct et la multiplication dans le domaine fréquentiel [5]. Cette propriété s'applique également du domaine fréquentiel vers le domaine direct, c'est-à-dire qu'une convolution dans le domaine fréquentiel correspond à une multiplication dans le domaine direct [5].

## Propriété de différentiation

La transformée de Fourier de la dérivée d'une fonction s'exprime comme la transformée de la fonction multipliée par un terme $-i\omega$ [76] :

$$\mathcal{F}\left\{\frac{\partial f(t)}{\partial t}\right\} = \frac{\partial}{\partial t}\mathcal{F}\{f(t)\} = -i\omega f(\omega). \tag{A.7}$$

Par conséquent, la *TF* de la dérivée d'ordre $n$ d'une fonction $f$ fait apparaitre un terme $(-i\omega)^n$ dans le domaine de Fourier.

## Propriété de glissement

Une multiplication d'une fonction $f$ par une exponentielle complexe correspond à une translation de la *TF* du signal dans le domaine réciproque [76]. Par conséquent, dans le domaine direct, on a :

$$f(t) \cdot e^{i\omega_0 t} = f(\omega - \omega_0), \tag{A.8}$$

et similairement dans le domaine de Fourier :

$$f(\omega) \cdot e^{i\omega t_0} = f(t - t_0), \tag{A.9}$$

où $t_0$ et $\omega_0$ sont des constantes.

### Propriété de conservation d'énergie (théorème de Parseval)

Le théorème de Parseval exprime l'équivalence des intégrales du produit d'une fonction et du complexe conjugué d'une autre exprimées dans des domaines réciproques [14]:

$$\int_{-\infty}^{\infty} f(t) \cdot \overline{g(t)}\, dt = \int_{-\infty}^{\infty} f(\omega) \cdot \overline{g(\omega)}\, d\omega, \tag{A.10}$$

où la notation $\overline{g(x)}$ indique le complexe conjugué de $g(x)$.

De l'Éq. (A.10), si $f(t) = g(t)$ :

$$\int_{-\infty}^{\infty} |f(t)|^2\, dt = \int_{-\infty}^{\infty} |f(\omega)|^2\, d\omega. \tag{A.11}$$

On peut interpréter l'équation (A.11) comme la « conservation d'énergie » entre le domaine direct et le domaine spectral lorsque l'énergie est proportionnelle au carré d'une quantité mesurable $f(t)$, comme dans le cas de la pression acoustique ou la différence de potentiel.

## A.2 L'équation d'onde

Soit un milieu gazeux non-visqueux, homogène et isotrope. On considère un petit élément cubique du milieu gazeux de volume $\Delta x \Delta y \Delta z$ illustré à la Figure A.1. Des forces sont appliquées dans la direction $x$ sur les surfaces d'aire $\Delta y \Delta z$.

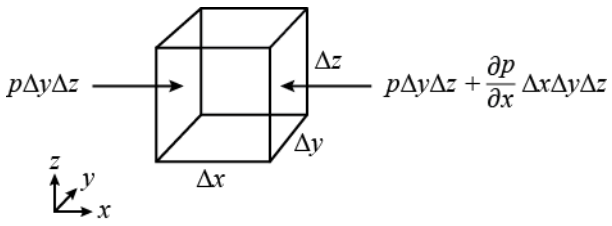


*Figure A.1 : Forces agissant sur un élément volumique élémentaire du fluide*

Puisque la pression est la force par unité d'aire, on exprime la force appliquée sur la surface de gauche comme $p\Delta y\Delta z$, et similairement sur la surface de droite en plus d'un terme variationnel $p\Delta y\Delta z + \frac{\partial p}{\partial x}\Delta x\Delta y\Delta z$ où $p(x, y, z, t)$ est la pression. La variation de force nette dans la direction $x$ s'exprime donc comme :

$$\Delta F_x = -\frac{\partial p}{\partial x}\Delta x\Delta y\Delta z. \tag{A.12}$$

En appliquant la même procédure dans les directions $y$ et $z$, on détermine la variation de force nette sur le volume élémentaire en fonction de la pression :

$$\Delta F = -\nabla p\Delta x\Delta y\Delta z, \tag{A.13}$$

où l'opérateur ∇ représente le gradient.

On établit également une relation entre la variation de la force et la variation du déplacement du fluide $\rho(x, y, z, t)$ en appliquant la deuxième loi de Newton :

$$\Delta F = \rho_0\Delta x\Delta y\Delta z\frac{\partial^2\rho}{\partial t^2}, \tag{A.14}$$

où $\rho_0$ est la densité volumique du fluide.

En égalisant les équations (A.13) et (A.14), et en considérant la vélocité particulaire $v = \frac{\partial\rho}{\partial t}$, on obtient l'équation d'Euler [13] qui met en relation $p$ et $v$ :

$$\nabla p = -\rho_0\frac{\partial v}{\partial t}. \tag{A.15}$$

On établit une autre relation entre $p$ et $\rho$ en considérant la dilatation volumique du volume élémentaire. Selon la loi de Hook, on sait que la dilatation varie linéairement avec la force (ou la pression) exercée [4], ce qui permet d'écrire :

$$p = -B\,\nabla\cdot\rho, \tag{A.16}$$

où ∇ · est l'opérateur divergence et $B$ est le module d'élasticité isostatique [11].

Finalement, on obtient l'équation d'onde en prenant la divergence de l'équation d'Euler (A.15) et en substituant la divergence du déplacement par la pression :

$$\nabla^2 p(x,y,z,t) - \frac{1}{c^2}\frac{\partial^2 p(x,y,z,t)}{\partial t^2} = 0. \tag{A.17}$$

L'opérateur Laplacien $\nabla^2 = (\nabla \cdot \nabla)$ est défini comme la divergence du gradient et $c = \sqrt{B/\rho_0}$ est la vitesse du son dans le fluide. On suppose que le fluide est non-dispersif, donc que la vitesse du son ne varie pas avec la fréquence de l'onde $\left(\frac{\omega}{k} = \frac{\partial\omega}{\partial k} = c\right)$.

De manière équivalente, l'équation d'onde s'exprime dans le domaine des fréquences $\omega$ en y appliquant une transformée de Fourier temporelle. Dans le domaine de Fourier, l'opérateur différentiel $\frac{\partial^2}{\partial t^2}$ fait apparaitre le terme $(-i\omega)^2 = -\omega^2$ tel que discuté à la sous-section 0. On obtient alors l'équation d'Helmholtz [2, 13]:

$$\nabla^2 p(x,y,z,\omega) + k^2 p(x,y,z,\omega) = 0. \tag{A.18}$$

## A.3 Théorème du gradient

Ce théorème, un cas particulier du théorème de Stokes [76], est analogue au théorème du flux-divergent mais s'applique à un champ scalaire tel que la pression. Il met en relation les intégrales volumique et surfacique définies par une surface fermée $S$ de volume $V$ englobant une source $A$ qui rayonne un champ scalaire. La géométrie est illustrée à la Figure A.2.

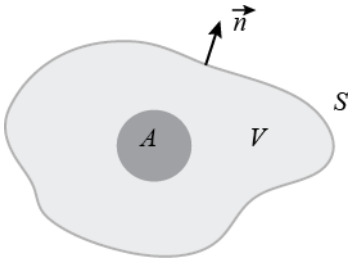


*Figure A.2 : Géométrie du théorème du gradient*

Le théorème du gradient s'exprime comme suit :

$$\iiint_V \nabla p \, dV = \oiint_S p \, d\vec{S}, \tag{A.19}$$

où $p$ est un champ scalaire, $d\vec{S} = \vec{n} \cdot dS$ et $\vec{n}$ est un vecteur unitaire orthogonal à la surface $dS$. Le théorème reflète une loi de conservation : l'intégration de la densité du champ est équivalente au flux sortant. Cette relation est fondamentale dans la résolution de problèmes de champs et est à la base de la méthode de NAH.

## A.4 Seconde identité de Green

L'intuition derrière le formalisme de Green est que le théorème du gradient reste valide même si le champ considéré n'est pas une fonction physique. En ce sens, on peut par exemple multiplier le champ physique $p$ par n'importe quelle fonction $g$ dérivable et intégrable, et appliquer l'équation (A.19) à ce produit :

$$\iiint_V \nabla(g \cdot p) \, dV = \iint_S g \cdot p \, d\vec{S}. \tag{A.20}$$

Le produit $(p \cdot g)$ n'a donc pas d'interprétation physique évidente. L'astuce mathématique est cependant de considérer la fonction $(p\nabla g - g\nabla p)$. Cette astuce nécessite que les dérivés simples et doubles de $g$ et $p$ soient continues. En insérant cette fonction dans (A.19), on obtient :

$$\iiint_V \nabla(p\nabla g - g\nabla p) \, dV = \iint_S p\nabla g - g\nabla p \, d\vec{S}, \tag{A.21}$$

qui se simplifie en :

$$\iiint_V (p\nabla^2 g - g\nabla^2 p) \, dV = \iint_S p\frac{\partial g}{\partial n} - g\frac{\partial p}{\partial n} \, dS, \tag{A.22}$$

où $\frac{\partial \mathrm{g}}{\partial n} = \nabla \mathrm{g} \cdot \vec{n}$ et idem pour $p$. L'équation (A.22) correspond à la seconde identité de Green [76].

## Fonction de Green

Il reste maintenant à choisir une fonction $g$ adéquate. La méthode de Green permet de résoudre les problèmes linéaires. Ceux-ci ont la forme :

$$Lu(x) = f(x), \tag{A.23}$$

où $L$ est un opérateur différentiel linéaire, $u(x)$ est la solution recherchée et $f(x)$ est le terme hétérogène. L'objectif de la méthode est d'obtenir une solution sous la forme intégrable :

$$u(x) = \int_{-\infty}^{\infty} g(x - x') \cdot f(x')\, dx'. \tag{A.24}$$

On reconnait un produit de convolution entre le terme hétérogène et la fonction de Green $g(x, x')$, qui est la réponse impulsionnelle du système. En quelque sorte, on peut interpréter $g$ comme l'opérateur inverse de $L$. On obtient $g$ en résolvant l'équation différentielle :

$$Lg(x, x') = \delta(x - x'). \tag{A.25}$$

La forme de la fonction de Green $g$ est indépendante de la valeur de $x$, c'est-à-dire que la forme de $G$ est propre à l'équation différentielle et est la même, peu importe le point où l'on cherche à résoudre l'équation différentielle. On peut retrouver l'équation (A.24) en multipliant l'équation (A.25) par $f(x')$ et en intégrant par rapport à $x'$ :

$$\int_{-\infty}^{\infty} LG(x, x') f(x')\, dx' = \int_{-\infty}^{\infty} f(x')\delta(x - x')\, dx' = f(x). \tag{A.26}$$

Selon l'équation (A.23), nous avons que :

$$Lu(x) = L \int_{-\infty}^{\infty} G(x, x') f(x')\, dx' \tag{A.27}$$

et puisque $L$ est un opérateur linéaire, on peut l'annuler de part et d'autre de l'équation (A.27) et retrouver ainsi l'équation (A.24).

## A.5 Formulation de Green pour l'équation d'Helmholtz

On applique donc le formalisme de Green à l'équation d'onde (A.17) afin d'obtenir la formulation analytique de la méthode d'holographie acoustique. On choisit d'abord de traiter le problème à partir de l'équation d'Helmholtz (A.18), puisque l'opérateur différentiel ne dépend que des variables d'espace. Selon l'équation (A.25), l'équation de Green associée est:

$$\nabla^2 g(r,\omega) + k^2 g(r,\omega) = \delta(r - r'), \tag{A.28}$$

avec $r = \sqrt{x^2 + y^2 + z^2}$.

En insérant (A.28) dans la seconde identité de Green défini à l'équation (A.22), le terme de gauche est simplifié, puisque $\nabla^2 p = -k^2 p$ et $\nabla^2 g(r,\omega) = -k^2 g(r,\omega) + \delta(r - r_0)$. On obtient alors :

$$p(r') = \iint_S p\frac{\partial g}{\partial n} - g\frac{\partial p}{\partial n}\, dS. \tag{A.29}$$

La solution de l'équation (A.28) est [4, 77, 14]:

$$g(r,r',\omega) = \frac{\mathrm{e}^{i\frac{\omega}{c}\sqrt{(x-x\prime)^2+(y-y\prime)^2+(z-z\prime)^2}}}{\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}} = \frac{e^{ikR}}{R}, \tag{A.30}$$

avec $R = |r - r'|$. Cependant, afin de simplifier (A.29) d'avantage, l'astuce mathématique est de considérer la solution de Dirichlet exprimée comme :

$$g_D(r,r_0) = g(r,r') - g(r,r_i), \tag{A.31}$$

avec $r_i = x' + y' + (2z_0 - z')$, qui est également la solution de (A.28).

# ANNEXE B ARTICLE 4 : METHOD FOR FINDING OPTIMAL EXPONENTIAL DECAY COEFFICIENT IN NUMERICAL LAPLACE TRANSFORM FOR APPLICATION TO LINEAR CONVOLUTION

J.-M. Attendu, A. Ross, publié dans *Signal Processing*, 130, pp. 47-56, 2017.

## B.1 Préface

Initialement, les principales idées derrière cet article ont été développées pour faire suite aux travaux de Blais et Ross qui formalisent l'holographie acoustique temporelle avec la transformée de Laplace [46]. Les principales motivations sont de déterminer l'effet de la NLT sur les fuites spectrales et d'offrir une alternative à l'application du *zero-padding* pour le calcul de la convolution linéaire. En effet, pour chaque dimension, le *zero-padding* augmente d'un facteur deux le nombre de données à traiter ce qui accroit les coûts en calculs, mais surtout en mémoire. Ainsi, avec une même capacité en mémoire, la transformée de Laplace permet le calcul de la convolution linéaire en considérant un fenêtrage deux fois plus large pour chaque dimension, ce qui permet de réduire les erreurs de troncature en NAH. La convolution calculée avec la transformée de Laplace permet d'obtenir la moitié du résultat de la convolution linéaire. Dans le domaine temporel, puisque le champ converge rapidement vers zéro, la première moitié est souvent suffisante, et ainsi, le calcul de la seconde moitié du champ propagé est évité.

Étant donné que l'application de la convolution avec la NLT s'étend à un cadre plus large que seulement le NAH, cet article a été développé dans ce sens et a été publié dans un journal de traitement du signal.

## B.2 Abstract

In this paper, a method based on the Numerical Laplace Transform is used for calculating the full linear convolution of real or complex signals. An algorithm for obtaining the last $N$ values of the convolution is presented, along with a method for finding an optimal value for the decay coefficient of the transform. It is shown that the use of the numerical Laplace transform formulation allows the calculation of each half of the linear convolution independently, which has computational benefits. The numerical Laplace transform is expressed as the fast Fourier transform of signals that

have been pre multiplied by a decreasing exponential window characterized by decay coefficient *c*. The error of the resulting linear convolution depends on the value of the decay coefficient; undervalue results in the generation of wrap-around error whereas overvalue causes amplification of Gibbs phenomenon. In this paper, a formula that optimizes the value of the decay coefficient is developed. A trade-off value for *c* is obtained and error analysis shows that it outperforms other coefficients proposed in the literature when applied to the calculation of linear convolution. The relative errors obtained are of the order of $10^{-6}$ % and $10^{-9}$ % for single and double precision.

## B.3 Introduction

The numerical Laplace transform (NLT) was introduced by Wilcox [78] for the analysis of linear systems. It has been used for such motives by others since, notably in the analysis of transients [79, 80, 81, 82]. In recent years, a renewed interest has occurred towards the NLT for its ability to improve the computation of linear convolutions.

Linear convolution is usually calculated by applying circular convolution on signals doubled with zero-padding [22]. However, this method increases significantly the amount of data to process, and consequently the computational complexity, especially in the case of 2D or 3D convolutions. Algorithms using the generalized discrete Fourier transform (GDFT) were used in order to obtain full linear convolution at reduced computational cost [20, 21]. Babic and Mandic proposed a similar method for obtaining full linear convolution with one circular convolution, by embedding the first non-wrapped samples in the high-range of the full dynamic range given by finite arithmetic precision, and the last *N* wrapped samples in the low-range, at the cost of sacrificing numerical accuracy [83]. The NLT, a specific case of the GDFT, was also used for such calculation by Martinez [20], but only to obtain the first *N* convolution values.

In this paper, a method for obtaining full $2N - 1$ values of the linear convolution with the NLT is presented. The main asset of the NLT is that it allows a form of control over the periodical repetitions that occur due to the discrete nature of the signal. With this property, it is possible to suppress wrap-around error from circular convolutions. This allows the computation of linear convolution without doubling signals with zero-padding, which is computationally beneficial. With the NLT, it is also possible to obtain the first or last *N* values of the linear convolution with the

calculation of only one circular convolution, unlike the GDFT which requires two calculations when convolving signals with complex values.

A succinct review of the numerical Laplace transform is presented along with some properties and its application to linear convolution. The main benefits of using the NLT for such application are described, along with a method for calculating the full linear convolution of complex signals using the NLT. The core of this paper deals with the elaboration of a simple formula to evaluate the optimized exponential decay coefficient used in the NLT. This coefficient is a trade-off value between the reduction of the wrap-around error, errors due to computer machine and Gibbs phenomenon. Wilcox [78], Wedepohl [80] and Inoue [40] have also worked in order to find such trade-off value. However, their analysis was not aiming at optimizing its application to linear convolution (no such work was found in the literature); rather, it aimed at optimizing the representation of analytical functions in the Laplace domain. Since similar numerical phenomena are involved in both applications, our results are analyzed and compared to those obtained with their methods.

The proposed method is applied to the resolution of an impulsive excitation of a 1D vibrational system, and to the convolution of different simple mathematical signals to bring out the impact of Gibbs phenomenon.

## B.4 Numerical Laplace transform

### Relation to the generalized Fourier transform

The numerical Laplace Transform (NLT) is a special case of generalized discrete Fourier transform (GDFT). For discrete signal $x[n]$, $n = \{0, \dots, N-1\}$, GDFT is defined as [84]:

$$F_\alpha(x[n]) = X_\alpha[k] = \sum_{n=0}^{N-1} x[n] e^{i\alpha n/N} e^{-2\pi i n k/N}. \tag{B.1}$$

Its inverse is:

$$F_\alpha{}^{-1}(X_\alpha[k]) = x[n] = e^{-i\alpha n/N} \sum_{n=0}^{N-1} X_\alpha[k] e^{2\pi i n k/N}, \quad \text{(B.2)}$$

where $\alpha \in \mathbb{C}$ is the modulation coefficient, and $k = \{0, \dots, N-1\}$ is the frequency vector. Equation (B.1) is equivalent to the regular discrete Fourier transform (DFT) of signal $x[n]$ initially multiplied with a complex exponential. Similarly, (B.2) is analogous to the inverse DFT demodulated using the complex exponential of opposite sign. The standard DFT case is obtained with $\alpha = 0$.

As discussed by [20, 40], in the case of real signals, full 2*N*-1 values of linear convolution using GDFT with complex modulation can be obtained by performing only one circular convolution of length *N*, with specific modulation parameter $\alpha$ in (B.1). For example, with $\alpha = \pi/2$, concatenation of the real and imaginary parts of the resulting circular convolution produces the linear convolution. This is true only if the convolved signals are real. It results in an improvement of the computational complexity of about 50 % in comparison with standard zero-padding method. However, when convolving complex signals, such as when the impulse response of a physical system is complex, two circular convolutions are required to obtain the linear convolution.

## Mathematical formulation

NLT is obtained when the modulation coefficient is chosen to be positive and imaginary ($\alpha = i \cdot c, c \in \mathbb{R}, c > 0$). Consequently, values of the complex exponential in (B.1) are real and decrease as *n* increases. For this reason, *c* is referred as the "damping" or "decay" coefficient [46] because multiplication of resulting decreasing exponential causes an attenuation of the signal with respect to time.

NLT, expressed with discrete Fourier transform (DFT), is defined as [85]:

$$NLT(x[n]) = X_c[k] = \sum_{n=0}^{N-1} \left(x[n] e^{-cn/N}\right) e^{-2\pi i n k/N} = DFT\left(x[n] e^{-cn/N}\right). \quad \text{(B.3)}$$

The inverse NLT is:

$$NLT^{-1}(X_c[k]) = x[n] = e^{cn/N} \sum_{n=0}^{N-1} X_c[k] e^{2\pi i n k/N} = e^{cn/N} DFT^{-1}(X_c[k]). \quad \text{(B.4)}$$

A main asset of the NLT (and the GDFT) is that by being very close to the DFT, it can be computed using fast Fourier transform (FFT) algorithms, making the process very efficient.

## Property 1: Reduction of the truncation error

Another asset of the NLT is that it attenuates truncation effects in the time domain in the case of any causal signal $x(t) = 0$ for $t < 0$. It thereby acts as a substitute to windowing. This property is brought up by expressing the discretization of the analytical signal and applying Laplace transform.

Multiplication of analog signal $x(t)$ with a Dirac comb $\Pi[n]$ is performed to obtain discretization. The Dirac combs in the time and frequency domains are expressed as:

$$\begin{aligned} \Pi(t) &= \sum_{n\in\mathbb{Z}} \delta(t - n\Delta t), \\ \Pi(\omega) &= \frac{2\pi}{\Delta t}\sum_{k\in\mathbb{Z}} \delta\left(\omega - \frac{2\pi k}{\Delta t}\right). \end{aligned} \tag{B.5}$$

By applying the convolution property of the Fourier transform, we obtain the following equation for $x(t)$:

$$\int_{-\infty}^{\infty} X(\omega - \omega') \cdot \Pi(\omega')\, d\omega' = 2\pi \int_{-\infty}^{\infty} x(t) e^{-\frac{ct}{T}} e^{-i\omega t} \Pi(\mathrm{t})\, dt. \tag{B.6}$$

Integrations are simplified by the Dirac combs, and equation (B.6) reduces to:

$$\frac{2\pi}{\Delta t}\sum_{k\in\mathbb{Z}} X\left(\omega - \frac{2\pi k}{\Delta t}\right) = 2\pi \sum_{n\in\mathbb{Z}} x(n\Delta t) e^{-\frac{cn}{N}} e^{-i\omega n\Delta t}. \tag{B.7}$$

As shown by Inoue [40], $X(\omega)$ can be isolated which brings out two error terms:

$$X(\omega) = \hat{X}(\omega) + E_1 + E_2. \tag{B.8}$$

Here, $\hat{X}(\omega)$ is analogous to the numerical Laplace transform of $x$:

$$\hat{X}(\omega) = \Delta t \sum_{n=0}^{N-1} x(n\Delta t) e^{-\frac{cn}{N}} e^{-i\omega n\Delta t}. \quad \text{(B.9)}$$

Factor $E_1$is the discretization error (or aliasing):

$$E_1(\omega) = -\sum_{k=1}^{\infty} \left[ X\left(\omega - \frac{2\pi k}{\Delta t}\right) + X\left(\omega + \frac{2\pi k}{\Delta t}\right) \right]. \quad \text{(B.10)}$$

The periodical repetitions of $X(\omega)$ brought out in (B.10) are therefore a consequence of discretization in the time domain. The repetitions are spaced proportionally to the sampling frequency.

In most engineering applications, function $X(\omega)$ has a spectral content that tends to zero for high frequency values. Thus, by taking a sufficiently high sampling frequency, $E_1$ can be neglected.

Factor $E_2$ is the truncation error:

$$E_2(\omega) = \sum_{n=N}^{\infty} x(n\Delta t) e^{-\frac{cn}{N}} e^{-i\omega n\Delta t}. \quad \text{(B.11)}$$

From (B.11), it is clear that a high value of *c* attenuates the effects of the truncation errors.

## Property 2: attenuation of periodic repetitions

Similarly, as a consequence of the application of inverse NLT over discretized signals, repetitions also occur in the time domain. As shown by the generalized Poisson summation introduced by Martinez, Heusdens and Hendriks [20], the $p^{\text{th}}$ periodic repetition of signal $x[n]$ is multiplied by $e^{-c \cdot p}$, where $p \in \mathbb{Z}$:

$$\sum_{p\in\mathbb{Z}} e^{-c \cdot p} x[n + pN] = \frac{e^{cn/N}}{N} \sum_{k=0}^{N-1} X_c[k] e^{2\pi i k n/N}. \quad \text{(B.12)}$$

The generalized Poisson summation formula links the time signal $x[n+pN]$ to the samples in the NLT spectrum. In other words, operations on $X_c[k]$ in the Laplace domain is equivalent to applying corresponding transformations in the time domain on $x[n+pN]$, considering periodical replications

modulated by $e^{-c \cdot p}$. Figure B.1 shows a simple example of signal $x[n+pN]$ with its modulated replications for $c = \ln(2)$ for $p$={-1,0,1,2}, which corresponds to multiplying $p^{th}$ repetitions by $2^{-p}$.

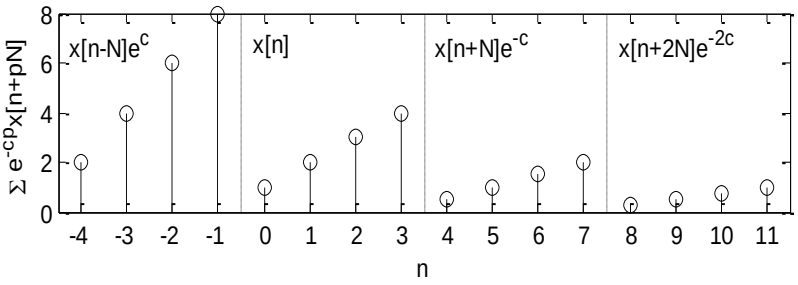


*Figure B.1 : Signal x[n] and its modulated periodical repetitions for c=ln(2)*

The modulation acts as a form of control on the periodic repetitions; by changing the value of damping coefficient $c$, it is possible to control the attenuation of the replicas. Applications of this property are numerous. For instance, Martinez used it to calculate acoustic auralization, using a damping coefficient corresponding to room absorption [84]. However, in many cases, the goal is to minimize the effect of the repetitions, and according to (B.12), this is obtained for $c \to \infty$. However, numerical errors generated in the frequency domain are multiplied by $e^{cn/N}$ in the application of the inverse NLT, resulting in the amplification of these errors. Consequently, if the value of $c$ is too high, these errors become considerable.

## Conditions of existence of the discrete-time Laplace transform

The conditions of existence of the discrete-time Laplace transform are closely related to those of the discrete-time Fourier transform (DTFT) [22], which is defined as:

$$X(\omega) = \sum_{n=-\infty}^{\infty} x[n] e^{-i\omega n}, \tag{B.13}$$

and its inverse is:

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega)\, e^{i\omega n} d\omega. \tag{B.14}$$

A signal $x[n]$ is causal if $x[n] = 0 \; \forall \; n < n_0$ and anti-causal if $x[n] = 0 \; \forall \; n > n_0$. According to Titchmarsh's theorem, any one of the following statements implies the other two [86, 87]:

1. $x[n]$ is either causal or anti-causal, and is of finite energy ($x[n] \in L^2$).
2. $X(\omega = \omega_R + i\omega_I)$ belongs to $L^2$ and is analytic in the complex plane for;
    a. $\omega_I > 0$ if $x[n]$ is causal,
    b. $\omega_I < 0$ if $x[n]$ is anti-causal and $X(\omega)$ approaches $X(\omega_R)$ as $\omega_I \to 0$.
3. The real and imaginary parts of $X(\omega)$ are the Hilbert transforms of each other.

If $x[n]$ is obtained experimentally, its physical measurement implies its causality and its finite energy, and we can confirm that the criteria above define the analyticity and finite energy of its Fourier transform.

The discrete-time Laplace transform is equivalent to the DTFT of $s[n] = x[n] \cdot e^{-cn/N}$, with $c \in \mathbb{R}$ and $c \geq 0$. The exponential window does not alter the causality nor the finite energy of $x[n]$, since $e^{-cn/N}$ monotonically decreases to zero. Therefore, we can conclude that the conditions on the analyticity and finite energy of the discrete-time Laplace transform are equivalent to those of the DTFT.

For the GDFT, similar relations to analyticity are obtained, as discussed by Martinez et al. [84].

## B.5 Linear convolution using the NLT

One of the main applications of the modulation property of NLT is to perform linear convolution. In fact, by using an appropriate value of *c*, the contributions of the repetitions are damped, which is a substitute to using zero-padding.

It is supposed that convolved signals *x*[*n*] and *y*[*n*] both have *N* values. Linear convolution is obtained by concatenating two circular convolutions of length *N*, resulting in a convolution of length 2*N-1* after discarding the value that is common to both circular convolutions. The first half of the linear convolution is expressed as:

$$NLT^{-1}\big(NLT(x[n])NLT(y[n])\big) = \sum_{m=0}^{n} x[m]y[n-m] + e^{-c}\sum_{m=n+1}^{N-1} x[m]y[N+n-m], \tag{B.15}$$

with $n \in \{0, \dots, N-1\}$. The last term corresponds to the wrap-around error, a consequence of periodical repetitions in the time domain. When $e^{-c} \approx 0$, wrap-around is suppressed and the first *N* values of linear convolution are obtained.

## Proposed method

In the present work, we have found that the last *N* values can be obtained by applying the same operation on reversed vectors *x*[*N*-*n*-1] and *y*[*N*-*n*-1] with $n \in \{0, \dots, N-1\}$, and by reversing the convolution result:

$$\begin{aligned} &NLT^{-1}\big(NLT(x[N-n-1])NLT(y[N-n-1])\big) \\ &= \sum_{m=0}^{n} x[N-m-1]y[N-n+m-1] \\ &+e^{-c}\sum_{m=n+1}^{N-1} x[N-m-1]y[m-n-1], \\ &n \in \{0, \dots, N-1\}. \end{aligned} \tag{B.16}$$

Concatenation of the result from (B.15) and the reversed result of (B.16) is applied to obtain the linear convolution of *x*[*n*] and *y*[*n*]. These operations are illustrated in Figure B.2, where symbol $*$ represents the NLT convolution operation:

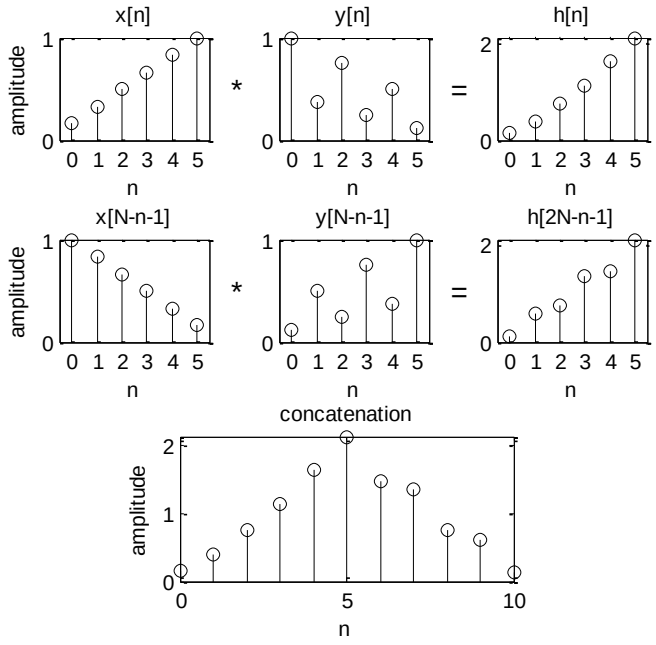


*Figure B.2 : Reconstruction of linear convolution of vectors* x*[*n*] and* y*[*n*] by concatenation of direct and reversed NLT convolution*

As shown in Figure 2, the last values of the two circular convolutions are repeated and thus, one must be discarded. A MATLAB code for the NLT convolution algorithm is provided in appendix A.

Note that there are other ways of obtaining the full linear convolution using weighted circular convolutions [20, 83, 84]. However, the method proposed here is applicable to real or complex signals, and each halves of the full linear convolution can be obtained without solving a two degrees linear system as it is required with the GDFT.

## Computational benefits

This analysis compares the computational efficiency of the present NLT method with the standard method to obtain linear convolution. The standard method consists in doubling the convolved signals with zero-padding before applying circular convolution with the FFT [22]. Operations involved in both processes are summarized by the block diagrams presented in Figure B.3. In the case of the NLT method in 3b, operations must be repeated twice to produce the full 2*N*-1 values of the linear convolution.

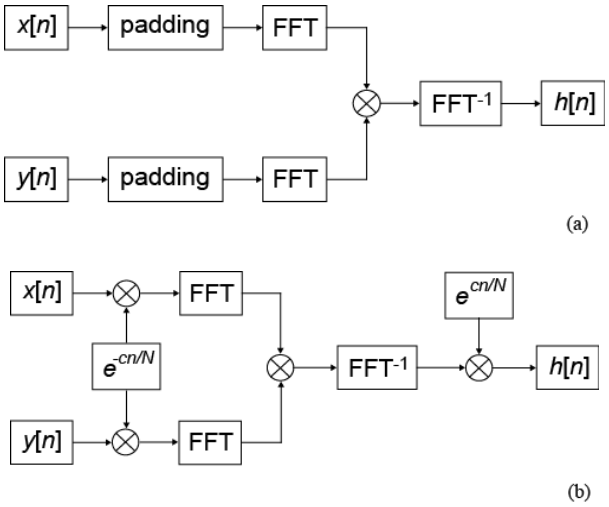


*Figure B.3 : Block diagrams of the operations involved in the calculation of the linear convolution with (a) the zero-padding method and b) the NLT method. The process shown in (b) produces half of the 2*N *values of the linear convolution*

Both methods require the use of the FFT, which is the most demanding operation. The performance analysis of the FFT is a complex matter because state-of-the-art FFT implementations are optimized for a large variety of cases. For example, the complexity gain can change with the presence of zero-pad or depending on the type of processor used. Consequently, the theoretical analysis presented in this paper is not absolute, and only gives a guideline for the evaluation of the complexity gains.

According to Franchetti et al., the fastest FFT libraries on multicore processors have a complexity with respect to input size $N$ in-between $4Nlog_2N$ and $5Nlog_2N$ floating-point operations (a conservative $5Nlog_2N$ is considered here) [88].

With the zero-padding method, by doubling $N$, the complexity becomes $10N \cdot \log_2[2N] = 10N \cdot \log_2[N] + 10N$ for each FFT. Thus, the three FFTs involved in the process require an order of complexity of $3 \cdot (10N \cdot \log_2[N] + 10N) = 30N \cdot \log_2[N] + 30N$. Additional $2N$ operations are considered for the multiplications in the Fourier domain, corresponding a total of $30N \cdot \log_2[N] + 32N$ operations.

In the case of the NLT method, the complexity of the FFT remains at $5N \cdot \log_2[N]$ since there is no padding. However, the process presented in Figure 3 (b) has to be performed twice to obtain the

first and the last $N$ values of the linear convolution. The application of the exponential windows requires $N$ multiplications for each application of the FFT. Also, 2N operations are added respectively for the multiplications in the Laplace domain and for reversing the order of the sequences. This results in a complexity of $2 \cdot 3 \cdot (5N \cdot \log_2[N] + N) + 4N = 30N \cdot \log_2[N] + 10N$. Following this analysis, although the complexity is slightly lower for the NLT algorithm, both methods necessitate similar computing requirements for high values of $N$.

Figure B.4 presents the calculation times for computing the linear convolution using the zero-padding method and the NLT method for the full $2N$-1 and $N$ values. Each calculation time is obtained from the average of 1000 computations. The calculations are performed on MATLAB using the FFT function, an adaptive FFT calculator based on the FFTW [89]. Zero-padding is added by using the embedded option of the function. An i5 Intel Quadcore processor (model 4690K, 3.5GHz) was used for the calculations.

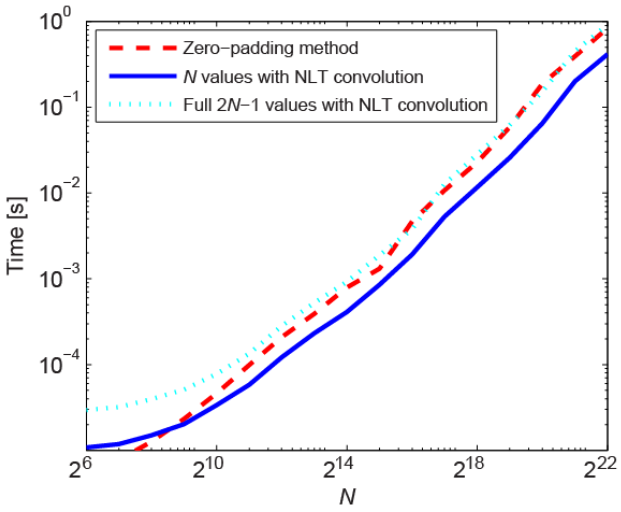


*Figure B.4 : Calculation times for the NLT convolution method and the zero-padding method with respect to the size of the convolved sequences*

According to Figure 4, our prediction for the performance of the algorithm is accurate, as both the zero-padding and full values NLT method give similar calculation times, and the computation of the $N$ values using the NLT represent a complexity gain of around 50 %.

One asset of the NLT algorithm is the fact that the linear convolution can be calculated in two distinct and independent steps, namely the first $N$ and the last $N$ convolution values. This has many advantages, the first being that if, for a certain application, only one half of the linear convolution

is required, this can be obtained with only 50 % of the calculations required by the zero-padding method. This can be the case, for example, in the analysis of transients; the behavior of such signal is often more interesting near its beginning. This is also an advantage of the NLT method over the GDFT method with $\alpha \in \mathbb{R}$, for which both circular convolutions are needed to obtain any part of the linear convolution when the convolved signals are complex.

Also, when convolving large signals or when performing multi-dimensional convolutions, a large amount of computer memory is required simultaneously. With the NLT method, only half the live memory is required.

## B.6 Optimization of the decay coefficient

Periodical repetitions in the time domain cause wrap-around error when performing numerical convolution using the NLT. According to (B.12), high values of damping coefficient $c$ attenuate the repetitions which results in decreased wrap-around error. In fact, when $c$ is high enough (near infinity), wrap-around is practically suppressed.

However, inverse NLT performed in the convolution process involves exponential amplification of the signal as shown in (B.4). Therefore, errors generated in the Fourier domain, including those caused by floating-point arithmetics, are increased exponentially through the application of the inverse NLT. For high values of $c$, the resulting error becomes considerable. Thus, there is a trade-off value for the damping coefficient that allows reducing wrap-around without significant amplification of the frequency domain error.

In this section, a review of existing methods for selecting $c$ is presented. Then, a new method is proposed and its application to numerical convolution is analysed.

### Review of existing methods

Some authors have proposed simple formulas to optimize the decay coefficient $c$. Two of these formulas are considered in this paper, since they appear to be the most discussed and the most used in the literature [40, 46, 85]. They were originally introduced by Wilcox [78] and Wedepohl [80] as empirical rules of thumb developed in order to approximate the analytical Laplace transform with the NLT. Therefore, they were not initially intended for optimizing the numerical Laplace convolution, even if they were applied to such operation previously [46].

Since similar numerical errors are generated in the inversion of the NLT regardless of the application, it is still relevant to compare the results obtained with Wilcox's and Wedepohl's coefficients to that of the proposed coefficient in the context of the calculation of linear convolution.

Wilcox proposed the following formula on the basis that it is simple and intuitively reasonable:

$$c = 2 \cdot \Delta\omega \cdot N = 2\pi \cdot f_s. \tag{B.17}$$

Wedepohl found heuristically that there is a relationship between the number of samples and an optimum choice of *c*:

$$c = \ln(N^2) \cdot \frac{N}{T} = 2\ln(N) \cdot f_s. \tag{B.18}$$

Here, (B.17) and (B.18) were adapted from their original forms in accordance with the definition of the NLT in (B.3) and (B.4).

Previous analysis by Moreno and Ramirez [85] shows that Wedepohl's criterion generally leads to more accurate results than Wilcox's criterion. Similar conclusions are obtained in our analysis and are discussed further in this paper.

## Proposed method

Wrap-around and amplification of frequency domain errors due to the NLT inversion, denoted respectively $E_W$ and $E_G$, are two independent phenomena that generate errors in the linear convolution obtained from the NLT. The goal is thus to find a trade-off value for the decay coefficient *c* that minimizes the sum of these errors:

$$min\{E_W(c) + E_G(c)\}, \tag{B.19}$$

It is important to understand that there is no efficient way to determine exact values for $E_W$ and $E_G$ without extensive calculations. Since an objective of the numerical Laplace convolution is to compute linear convolution at low machine costs, a simple means for evaluating these errors must be provided.

In section II, it is established that the time-domain repetitions, which are the cause of the wrap-around errors, decrease exponentially with $c$, and that the inverse NLT amplification increases exponentially with $c$. It can thus be assumed that $E_W(c)$ decreases exponentially with $c$ and that $E_G(c)$ increases exponentially with $c$. Analysis of the linear convolution error presented further in this paper confirms this exponential behavior. Thereby, the following model is posed:

$$E_W(c) = e^{-\gamma c} \cdot \varepsilon_W, \tag{B.20}$$

$$E_G(c) = e^{\beta c} \cdot \varepsilon_G, \tag{B.21}$$

where $\gamma$, $\beta$, $\varepsilon_W$ and $\varepsilon_G$ are constants. Equation (B.19) then simplifies to (B.22):

$$min\{e^{-\gamma c} \cdot \varepsilon_W + e^{\beta c} \cdot \varepsilon_G\}. \tag{B.22}$$

The minimum value of (B.22) is found by taking its derivative with respect to $c$:

$$\frac{\partial\left(e^{-\gamma c} \cdot \varepsilon_W + e^{\beta c} \cdot \varepsilon_G\right)}{\partial c} = -\gamma e^{-\gamma c} \cdot \varepsilon_W + \beta e^{\beta c} \cdot \varepsilon_G = 0. \tag{B.23}$$

The damping coefficient $c$ is isolated from (B.23), yielding:

$$c = ln\left(\frac{\varepsilon_W \cdot \gamma}{\varepsilon_G \cdot \beta}\right)^{\frac{1}{\gamma+\beta}}. \tag{B.24}$$

Now, the phenomena generating the errors are examined in order to find appropriate values for constants $\gamma$, $\beta$, $\varepsilon_W$ and $\varepsilon_G$.

**Wrap-around error**

From (B.15) and (B.16), it is shown that the wrap-around error $E_W$ generated by NLT convolution is a product of $e^{-c}$ with the resulting linear convolution $h[n]$:

$$\begin{aligned} E_W[n, c] &= e^{-c} \cdot h[n + N], \\ E_W[n + N, c] &= e^{-c} \cdot h[n], \\ n &\in \{0, \dots, N - 1\}. \end{aligned} \tag{B.25}$$

To fit the model proposed in (B.20), $\gamma$ must be set to one. The value of $\varepsilon_W$ is equal to the amplitude of the linear convolution in the time domain. Obviously, $h[n]$ cannot be used to find *c*, since the whole purpose of determining the optimal value of *c* is to calculate the linear convolution $h[n]$. Fortunately, $\varepsilon_G$ also depends on $h[n]$, and the dependency to $h[n]$ cancels after the substitution in (B.24).

**Inverse NLT error**

Inverse NLT error $E_G$ is a consequence of the amplification of frequency domain errors caused by the exponential term of the inverse NLT. From our analysis, we observe that the frequency domain errors are principally due to floating point arithmetic and Gibbs phenomenon. Other authors have also observed amplification of Gibbs errors due to NLT inversion [40, 90, 91].

In order to evaluate $E_G$, a heuristic approach is used. By observing the variation of the linear convolution error as a function of *c*, the influence of three main parameters on $E_G$ is noticed; the machine error $\epsilon$, the amplitude of $h[n]$ and the presence of discontinuities in $h[n]$. In fact, it is observed that $\varepsilon_G$ increases proportionally with the amplitude of $h[n]$ and the machine error $\epsilon$:

$$\varepsilon_G = \kappa \cdot \epsilon \cdot h[n] \tag{B.26}$$

where $\kappa$ is an unknown constant.

By inserting (B.26) in the model proposed in (B.21), we obtain:

$$E_G = e^{c\beta} \cdot \kappa \cdot \epsilon \cdot h[n] \tag{B.27}$$

From (B.4), it is known that the error is amplified by the exponential window $e^{cn/N}$. To fit the proposed model, the assumption that the error is evenly distributed over *n* is made and, consequently, we let the average of $e^{cn/N}$ be $e^{\beta c}$.

Based on this, the average of the exponential window can be determined since it is a geometric series:

$$\frac{1}{N}\sum_{n=0}^{N-1} e^{cn/N} = \frac{1}{N}\left(\frac{1-e^{c}}{1-e^{c/N}}\right). \tag{B.28}$$

Then, by comparing this value to the model:

$$\beta = \frac{1}{c} ln\left(\frac{1}{N}\left(\frac{1-e^{c}}{1-e^{c/N}}\right)\right), \tag{B.29}$$

it is found that the resulting function $\beta(N,c)$ has an asymptotic behavior for $N \gg c$. Since $N$ is large in most engineering applications, it is expected that the optimal value of $\beta$ does not depend on the length of the signal $N$. The asymptotic value of $\beta$, obtained from the analysis below, is taken for the model.

## B.7 Analysis

The objective of this section is to find adequate values for $\beta$ and $\kappa$, and to demonstrate the good performance of the proposed decay coefficient in the calculation of linear convolution. The analysis focuses on three parameters, namely, the length of the signals $N$, the numerical precision $\epsilon$, and the type of signals. The error is studied as a function of the damping coefficient $c$. First, the analysis of the auto-convolution of different simple mathematical functions is presented. This analysis leads to determining appropriate values for $\beta$ and $\kappa$. Then, a practical example of the application of the NLT convolution algorithm to predict vibration excitation of a simple system is presented. This analysis is performed for different signal lengths $N$.

It is shown that the proposed method gives satisfactory approximation of the optimal coefficient $c$ that minimizes the error. It also outperforms Wilcox's and Wedepohl's predictions for all cases.

### Method for evaluating the accuracy

In order to quantify the precision of the linear convolutions obtained with the proposed method, an analysis of the error is performed. The root-mean-square (RMS) deviation from the exact linear convolution $h_{ref}[n]$ is used to evaluate the overall accuracy, and is defined as:

$$E_{RMS} = \sqrt{\frac{\sum_{n=0}^{2N-1}\left(h[n] - h_{ref}[n]\right)^2}{2N-1}}. \tag{B.30}$$

The reference $h_{ref}[n]$ is obtained from the definition of linear convolution in its summation form, expressed in (B.31) and (B.32);

$$h_{ref}[n] = \sum_{m=0}^{n} x[m] \cdot y[n-m], \tag{B.31}$$

$$h_{ref}[n+N] = \sum_{m=0}^{n} x[m] \cdot y[N+n-m], \quad n \in \{0, \dots, N-1\}. \tag{B.32}$$

The relative error is the root-mean-square error $E_{RMS}$ normalized by the root-mean-square value of $h_{ref}$, and is denoted $e_{RMS}$.

## Error analysis of various simple functions and influence of the Gibbs phenomenon

Error analysis is performed for six simple signals of distinct shapes. Linear convolution of each function with itself (auto-convolution) is calculated with the NLT method. The different functions used are presented in Table 1, for $n \in \{0, \dots, N-1\}$.

*Tableau B.1: Mathematical description of the analyzed functions, and values of κ obtained from the simulations*

| *Functions* | *Mathematical expression* | *κ* |
|---|---|---|
| Rational | $\big((n+1)/N\big)^{-1}$ | 0.5 |
| Logarithmic | $\log((n+1)/N)$ | 0.06 |
| Quadratic | $(n/N)^2$ | 0.02 |
| Exponential | $e^{-2n/N}$ | 0.02 |
| Linear | $n/N$ | 0.01 |
| Constant | 1 | 0.005 |

Since every auto-convolution has different amplitude, the relative error $e_{RMS}$ is used as the comparative value. It is presented in Figure 5 as a function of the decay coefficient *c*. The normalized RMS error is presented on a logarithmic scale, while coefficient *c* is on a linear scale. As an example, the model predictions for the wrap-around error and inverse NLT error described in (B.25) and (B.27) are also presented as the dotted and dashed curves, with $\beta = 0.95$ and $\kappa = 0.05$. Note that the slopes of exponential functions in semi log10 scale are proportional to the slopes of the same functions in natural semi logarithmic scale. The error is shown for the double precision case, but similar results are obtained in single precision.

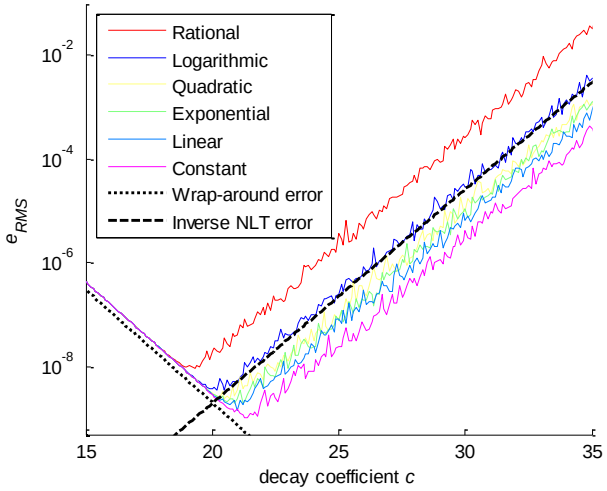


*Figure B.5 : Normalized RMS error obtained for the auto-convolution of six functions for N = 512 in double precision*

For low values of *c*, the error initially decreases identically for all signals. All test-signals have approximately the same normalized wrap-around error, since all curves superimpose on the left-hand side of Figure B.5. Also, it can be seen that the proposed model (B.20) with $\gamma = 1$ accurately predicts the behavior and the values of the wrap-around error.

Then, for larger values of *c*, the error increases exponentially with *c*, due to wrap-around and NLT inversion. From Figure 5, it can be seen that the error increases at similar exponential rates for all signal types, since the curves are parallel on the semi-logarithmic scale. According to (B.27), this means that the value of $\beta$ is constant regardless of the signal type. The slope corresponds to a value of $\beta = 0.95$. This value is thus applied to the model.

The optimal value $c_{opt}$ is found at the location of the minimum RMS error, that is, at the intersection of the wrap-around dominated and the NLT inversion dominated parts of the error curves. It varies for each test-signal between $c \approx 17$ to $22$. The auto-convolution yields different values for the inverse NLT error for each signal, as each curve is shifted vertically relative to each other. Therefore, the inversion error starts being significant at different values of $c$ in each case.

According to (B.27), the vertical shift of the inversion error is interpreted as a variation of coefficient $\kappa$ for the different signals. In table I, the values of $\kappa$ that fit the auto-convolution errors are presented. Coefficient $\kappa = 0.05$ appears to be an adequate trade-off value for the cases presented, as it can be seen from the dashed curve in Figure 5.

By pushing the analysis further, it is found that signals having the highest inversion error are the ones that are the most subject to Gibbs phenomenon. Gibbs phenomenon is the inability to recover point values of a non-periodic function from is Fourier coefficients [92]. The discrepancy is maximal around discontinuities in the time domain. By analysing the linear convolution results, signals that generate the highest errors (that have the highest $\kappa$ value) have discontinuities and no n-periodical behaviors.

In Figure 5, the rational signal has the highest inversion error. By taking a closer look at the linear auto-convolution of the rational signal, we notice the presence of manifest discontinuity. This is shown in Figure 6, where both the classical linear convolution (i.e. reference) and NLT linear convolution are shown near the discontinuity. With NLT convolution, oscillatory behavior is observed around the discontinuity, testifying of the presence of the Gibbs phenomenon. A value of $c = 35$ was used to exaggerate the amplification of the Gibbs phenomenon.

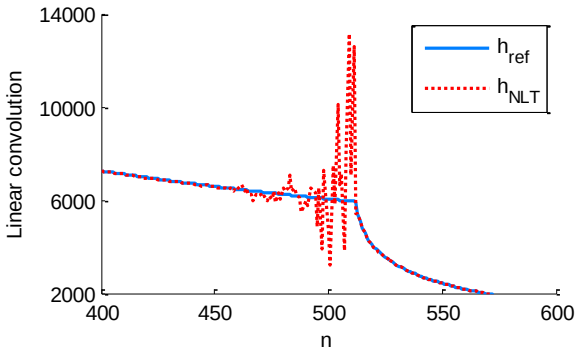


*Figure B.6 : Behavior near the discontinuity of the reference linear convolution and NLT linear convolution with $c = 35$ for the rational signal with N = 512 in double precision*

Thus, the amplitude of the Gibbs phenomenon error depends on the extent and number of discontinuities. It also depends on the length of the signals $N$, as the Gibbs phenomenon reduces as the number of points increases. Since the impact of the Gibbs phenomenon is not easily predictable, the value of $\kappa = 0.05$ is maintained since it is a good compromise for most signals tested by the authors.

Other authors have pin pointed the impact of the Gibbs phenomenon on the NLT inversion, and have proposed some methods to reduce its impact in the time domain, in particular by applying a filter in the frequency domain [90, 91].

## Optimal damping coefficient

By inserting (B.25) and (B.27) in (B.24), the following simplification is obtained:

$$c = \frac{1}{1+\beta} ln\left(\frac{1}{\kappa \cdot \epsilon \cdot \beta}\right) \tag{B.33}$$

By considering $\beta = 0.95$ and trade-off value $\kappa = 0.05$ in (B.33), the decay coefficient can be obtained. As an example, in MATLAB, $\epsilon \approx 1.2 \cdot 10^{-7}$ in single precision and $\epsilon \approx 2.2 \cdot 10^{-16}$ in double precision, which leads respectively to decay coefficients $c \approx 9.7$ and $c \approx 20.1$.

## Impulsive vibrational excitation of a 1D system

Linear convolution is performed to calculate the transient vibration response of a damped single-degree-of-freedom system subjected to impulse excitation [93]. Error obtained with the proposed coefficient is shown, along with errors obtained with Wilcox's and Wedepohl's coefficients. The force $x[n]$ is a short-time half sine impulse defined as:

$$x[n] = \begin{cases} \sin\left(\frac{16\pi n}{N}\right) & \text{for } n = \{0, \dots, N/16\}, \\ 0 & \text{for } n = \left\{\frac{N}{16} + 1, \dots, N-1\right\}. \end{cases} \tag{B.34}$$

The vibrational impulse response $y[n]$ is:

$$y[n] = \frac{1}{m\omega_d} e^{-\zeta n/N} sin(\omega_d n/N), \tag{B.35}$$

Where the physical constants are fixed at $m = 1\ kg$, $\omega_d = 12\pi$ rad/s and $\zeta = 2.5$.

The linear convolution of these signals is used to calculate the displacement of the system with respect to time. The two signals along with their convolution are illustrated in Figure 7.

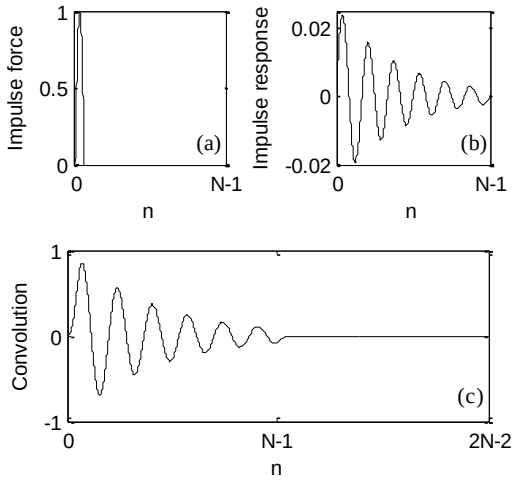


*Figure B.7 : (a) Force impulse* x[n]*, (b) impulse response* y[n] *and (c) linear convolution for N=1024*

The RMS error with respect to the damping coefficient is presented in Figure 8. The RMS error is given on a logarithmic scale, while *c* is on a linear scale. Results are presented both for single and double precision, and for different signal lengths.

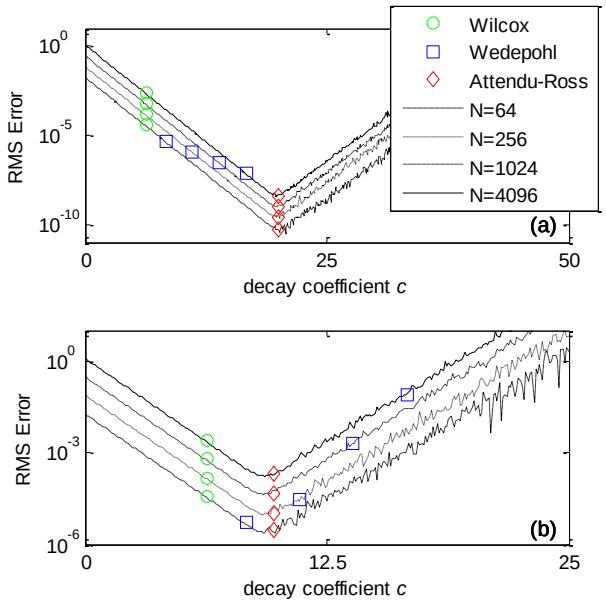


*Figure B.8 : RMS error obtained for the convolution of the force function with the impulse response in relation with the damping coefficient for* N = *64, 256, 1024, 4096 in (a) double precision and (b) single precision*

In Figure 8, two distinct behaviors are observed: for lower values of the decay coefficient, the error decreases with *c*, and for higher values, the error increases. In fact, these two behaviors are due to the wrap-around and inverse NLT errors, respectively, as presented previously. The total error is minimized at an optimum value $c_{opt}$. For this example, the damping coefficients found from eq. (B.31), represented by red diamonds on Figure 8, predict the optimal values within less than 5 % for both single and double precision.

The proposed method gives a better approximation of $c_{opt}$ than Wilcox's and Wedepohl's formulas, especially in the case of double precision. Wilcox's formula (green circles) under-estimates $c_{opt}$ by 33% and 69% in single and double precision, respectively. Wedepohl's results (blue squares) deviate from $c_{opt}$ in a range of 11% to 79% in single precision and 19% to 58% in double precision. Wedepohl's predictions are closer to the optimum in single precision for $N = 64$ and $N = 256$; however, the dependency to $N$ does not predict adequately the behavior of the error in this case.

For all cases shown in Figure 8, the optimal value remains approximately the same regardless of the length of the signals $N$. This is in conformity with the proposed model, since the supposes that $N \gg c$, as explained in section III.B with (B.29).

Figure 8 also shows that the RMS error increases with $N$. This is due to the amplitude of the linear convolution in (B.30) that also increases with $N$. By normalizing the error by the root-mean-square of $h[n]$, we obtain relative errors of $2.7 \cdot 10^{-2}$ % and $3.5 \cdot 10^{-6}$ % for single and double precision with $N = 4096$. For most engineering applications, errors of this magnitude are quite acceptable.

## B.8 Conclusion

In this paper, a method for obtaining full 2*N*-1 values of linear convolution using the FFT-based NLT is presented. With this method, the first and last halves of the convolution are calculated separately. In comparison with traditional zero-padding method, the NLT method requires only half the memory simultaneously during computation. Also, the proposed method can take advantage of parallel computation since the calculations of both halves of the linear convolution result are independent.

A formula to determine a trade-off value for the decay coefficient required by the NLT is presented. This formula is based on analytic derivations and confirmed by heuristic observations. It depends on four parameters: the machine error $\epsilon$, coefficient $\gamma = 1$ related to the wrap-around error, and coefficients $\beta = 0.95$ and $\kappa$, related to the NLT inversion error. The value of $\kappa$ depends on the susceptibility of the convolved signals to generate Gibbs error, and should be between 0.005 and 0.5. A value of $\kappa$=0.05 is shown to be a reasonable trade-off, yielding decay coefficient values of 9.7 and 20.1 in single and double precision, respectively. It is found that the machine error and the presence of discontinuities in the convolution signal are two main factors that contribute to the error. On the contrary, the signal length is not a significant factor.

Application of this NLT convolution to various signals is presented and analysis. The good performance of the method is shown; the relative error obtained in the case of a simple 1D vibrational system is of the order of $10^{-2}$ % and $10^{-6}$ % in single and double precisions, respectively.

Prospective work could include a more elaborated analysis for finding the optimal decay coefficient. Such analysis could rely less on heuristic observations by using a method like convex

optimization. Also, an adaptation of this method could be developed for multi-dimensional convolutions.

## B.9 Appendix A

A MATLAB code for computing the full 2*N*-1 values linear convolution with the NLT is provided:

```
function [h_nlt]=nlt_conv(f,g,c)
% Performs linear convolution using the
% NLT algorithm.
% f: signal #1.
% g: signal #2.
% c: damping coefficient, c>0.
% h_nlt: resulting linear convolution.

N = length(f); % f and g must have the same length
n=0:N-1;
h_nlt = inlt(nlt(f(end:-1:1),c).*nlt(g(end:-
1:1),c),c);
h_nlt = [inlt(nlt(f,c).*nlt(g,c),c) h_nlt(end-1:-
1:1)];
end


function [X]=nlt(x,c)
% Performs numerical Laplace transform.
% x: signal in the time domain.
% c: damping coefficient.
% X: signal in the frequency domain.


N=length(x);
n=0:N-1;
X=fft(exp(-c.*n./N).*x);
end


function [x]=inlt(X,c)
% Performs inverse numerical Laplace transform.
% X: signal in the frequency domain.
% c: damping coefficient.
% x: signal in the time domain.


N=length(X);
n=0:N-1;
x=exp(c.*n./N).*ifft(X);
end
```

# ANNEXE C ARTICLE 5: ON THE CHOICE OF THE MODULATION COEFFICIENTS FOR IMPROVED COMPUTATIONAL EFFICIENCY IN MULTI-DIMENSIONAL LINEAR CONVOLUTION USING THE GENERALIZED DISCRETE FOURIER TRANSFORM

J.-M. Attendu, A. Ross, soumis à *Signal Processing* le 9 juin 2017.

## C.1 Préface

Cet article présente un algorithme permettant de calculer plus efficacement la convolution linéaire multidimensionnelle. Son application à l'holographie acoustique temporelle est intéressante, puisqu'elle présente une alternative au calcul de la convolution 3-D avec la méthode du *zero-padding* qui est coûteuse en opérations et en mémoire. En effet, les calculs de propagation présentés aux articles 1 et 2 ont été effectués avec l'algorithme présenté à l'article 5. Cet article a été soumis à un journal de traitement du signal, puisque son application ne se limite pas qu'au contexte de l'holographie acoustique.

## C.2 Abstract

The generalized discrete Fourier transform is an efficient tool for the calculation of linear convolution. It was evaluated that for real signals, such algorithm can reduce the computational complexity by 50 % in comparison to standard zero-padding method. This paper shows that for multi-dimensional convolution, it is possible to reduce the complexity additionally by adequately choosing the modulation coefficients required by the generalized Fourier transform. In fact, this reduces the number of Fourier transform calculations required by the algorithm. By evaluating the floating point operations involved in the calculation, we show that the proposed algorithm leads to an additional reduction of the computational cost of about 10 % to 25 %, depending on the convolved signals. This reduction is verified experimentally by comparing calculation times obtained with the proposed algorithm and the standard method. Sets of coefficients that produce such reduction are presented for 2-D and 3-D cases. It is shown that the improvement in computational costs increases with the number of dimensions, and that the proposed method is also beneficial to the convolution of complex signals.

## C.3 Introduction

Standard linear convolution algorithms use zero-padding to suppress wrap-around error from circular convolution [22]. It requires doubling the data with zero values, which results in a considerable increase of operations and memory usage. The increase is even more significant with multi-dimensional convolution, since data is doubled with respect to each dimension. This paper proposes a method for reducing the computational complexity for applications in multi-dimensional linear convolutions. It is based on the use of weighted circular convolutions, which are obtained by multiplying the convolved signals with a complex exponential window, prior to calculating circular convolution with the fast Fourier transform (FFT) [83, 20, 84, 21]. The consequence of such modulation is that the periodical replications due to discretization in the frequency domain are multiplied by a constant that depends on the coefficient used in the exponential window. This allows the isolation of wrap-around caused by the replications from the linear convolution result. Others have shown that in the case of real unidimensional signals, the weighted circular convolution method allows reducing computational complexity to about half the complexity required by standard method [83, 20, 21]. Similar results were obtained using other number theoretic transforms, such as the Fermat number transform [94, 95]. An adaptation for the 2-D linear convolution was proposed by Radhakrishnan and Jenkins, which led to similar performances in the case of real signals [96, 50]. These results suppose that the efficiency of the FFT is the same for real and complex input signals, which is not the case in practice. This factor is considered in the present study.

In this paper, it is shown that the complexity can be further reduced in the multi-dimensional case (two dimensions or greater). This improvement comes from the fact that the calculation of the multi-dimensional linear convolution requires the combination of several weighted circular convolutions. In this process, several FFT computations can be spared by using the same weighting coefficients in the calculation of different circular convolutions.

An evaluation of the floating point operations required by the proposed method, along with that of the zero-padding and standard generalized discrete Fourier transform (GDFT) methods are presented for an arbitrary number of dimensions. These predictions are verified by comparing calculation times. The proposed method can contribute to numerous multi-dimensional linear convolution applications, such as image/video processing, multi-dimensional array filtering,

correlation, boundary value problems, etc. For example, the authors have successfully applied the proposed method to 3-D linear convolutions in time-domain near-field acoustical holography [54].

This paper is organized as follows: first, an overview of the multi-dimensional GDFT is presented, along with its application to the weighted and linear convolutions. Then, the method for choosing the modulation coefficients in order to improve the efficiency is presented, and modulation coefficient values for the two-dimensional and three-dimensional cases are proposed. This is followed by a computational complexity analysis that compares the proposed algorithm with standard zero-padding method. The results are developed for the multi-dimensional case, with an emphasis on the more common 2-D and 3-D cases. Finally, the theoretical predictions of the reduction of floating point operations are verified by comparing actual calculation times.

## C.4 The multi-dimensional GDFT and its application to linear convolution

In this section, an overview of the multi-dimensional GDFT is presented, along with its application to convolution. $M$-dimensional signals $x[n_1, \dots, n_M]$ are considered. For each dimension, the discrete finite domain is characterized by $n_i \in \{0,1, \dots, N_i - 1\}$. Vectors $k_1, \dots, k_M$ form the reciprocal frequency domains, with $k_i \in \{0,1, \dots, N_i - 1\}$. Throughout this paper, the direct and reciprocal domains are designated as *n*-space and *k*-space.

### Multi-dimensional GDFT

The multi-dimensional generalized discrete Fourier transform (MD-GDFT) of signal $x[n_1, \dots, n_M]$ is expressed as:

$$X_\alpha = \sum_{n_1,\dots,n_M=0}^{N_1-1,\dots,N_M-1} \Omega_\alpha \cdot x \cdot e^{-2\pi i\left(k_1\frac{n_1}{N_1}+\cdots+k_M\frac{n_M}{N_M}\right)}, \tag{C.1}$$

$$\Omega_\alpha = e^{i\left(\alpha_1\frac{n_1}{N_1}+\cdots+\alpha_M\frac{n_M}{N_M}\right)}.$$

The modulation window $\Omega_\alpha$ is defined by $M$ modulation coefficients $\alpha_1, \dots, \alpha_M$, each of which is related to one dimension. The MD-GDFT is equivalent to the $M$-dimensional DFT (or FFT) applied

on a signal multiplied by the exponential window $\Omega_\alpha$. The standard discrete Fourier transform (DFT) case is thus obtained when the modulation coefficients $\alpha_i$ are all zero. In 1-D, the skew-cyclic case discussed in the literature is obtained when $\alpha = \pi$ [21]. When $\alpha$ is imaginary, the modulation term is a decreasing exponential and (C.1) becomes equivalent to the numerical Laplace transform [47].

The inverse MD-GDFT applied to $X_\alpha$ is:

$$x_\alpha = \frac{{\Omega_\alpha}^{-1}}{N} \sum_{k_1,\dots,k_M=0}^{N_1-1,\dots,N_M-1} X_\alpha \cdot e^{2\pi i\left(k_1\frac{n_1}{N_1}+\cdots+k_M\frac{n_M}{N_M}\right)},$$
$$(C.2)$$
$${\Omega_\alpha}^{-1} = e^{-i\left(\alpha_1\frac{n_1}{N_1}+\cdots+\alpha_M\frac{n_M}{N_M}\right)},$$
$$\mathrm{N} = N_1 \cdot N_2 \cdot \ldots \cdot N_M.$$

The inverse modulation ${\Omega_\alpha}^{-1}$ is independent of $k$, and can be applied subsequently to the inverse Fourier transform.

As with the DFT, discretization in the *k*-space causes the signal $x_\alpha$ to be periodically repeated in the *n*-space [84]. With the GDFT, the $p^{th}$ periodical repetition is multiplied by constant $e^{i\alpha\cdot p}$, where $p \in \mathbb{Z}$. This is shown by the generalized Poisson summation formula derived by Martinez, Heusdens and Hendriks [20]. In the multi-dimensional case, similar phenomenon occurs: the $p^{th}$ periodical repetition with respect to each dimension is multiplied by geometrically weighted constant, as expressed in (C.3).

$$x_\alpha = \sum_{\substack{p_1\in\mathbb{Z}\\ \vdots \\ p_M\in\mathbb{Z}}} e^{i(\alpha_1\cdot p_1+\cdots+\alpha_M\cdot p_M)} \cdot x(n_1 + p_1N_1, \cdots, n_M + p_MN_M) \qquad (C.3)$$

## Weighted circular convolution

Weighted circular convolution $h_\alpha$ is obtained by multiplying the convolved signals in the *k*-space, and by applying the inverse MD-GDFT, as summarized in the block diagram presented in Fig. 1.

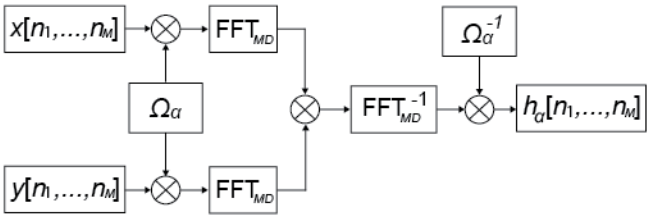


*Figure C.1 : Block diagram of the weighted circular convolution obtained with the generalized discrete Fourier transform*

Multiplying the periodical repetitions of the signals with the exponential constants has the following consequence to the weighted convolution result: the last $N$ values of the $2N$ values of the linear convolution (i.e. the wrapped-around terms, which are usually "folded" in circular convolution), are multiplied by the constants. Consequently, a mathematical relation is established between linear convolution $h$ and weighted circular convolution $h_\alpha$.

For example, it was shown that in 1-D, such relation is [84]:

$$h_\alpha[n] = h[n] + e^{i\alpha} h[n+N], \tag{C.4}$$

with $n \in \{0,1,\dots,N-1\}$. Here, the weighted circular convolution $h_\alpha$ has $N$ values and the linear convolution $h$ has $2N$ values.

Equivalently, it was shown by Radhakrishnan and Jenkins that the 2-D weighted circular convolution can be expressed as [50]:

$$\begin{aligned} h_\alpha[n_1,n_2] = h[n_1,n_2] + e^{i\alpha_1} h[n_1+N_1,n_2] \; + e^{i\alpha_2} h[n_1,n_2+N_2] \\ + e^{i(\alpha_1+\alpha_2)} h[n_1+N_1,n_2+N_2] \end{aligned} \tag{C.5}$$

The 3-D case is developed as follows:

$$\begin{aligned} h_\alpha[n_1,n_2,n_3] = h_{000} + e^{i\alpha_1}\cdot h_{100} + e^{i\alpha_2}\cdot h_{010} + e^{i\alpha_3}\cdot h_{001} + e^{i(\alpha_1+\alpha_2)}\cdot h_{110} + \\ e^{i(\alpha_1+\alpha_3)}\cdot h_{101} + e^{i(\alpha_2+\alpha_3)}\cdot h_{011} + e^{i(\alpha_1+\alpha_2+\alpha_3)}\cdot h_{111}, \end{aligned} \tag{C.6}$$

with the following notation:

$$h_{pqr} = h[n_1 + pN_1, n_2 + qN_2, n_3 + rN_3]. \quad \text{(C.7)}$$

Such formulation can be extended to other multi-dimensional cases; the weighted circular convolution of $M$-dimension signals results in the sum of $2^M$ terms.

## Linear convolution

These relations can be used to determine the linear convolution $h[n_1, \cdots, n_M]$ from $2^M$ different weighted circular convolution results $h_\alpha$. According to equations (C.4) to (C.6), the $2^M$ weighted circular convolutions form a linear system composed of the $2^M$ unknown linear convolution subarrays. In this system, each unknown subarray is multiplied by a constant that depends on the modulation coefficients of the weighted circular convolutions. The subarrays are determined by solving the system, and they are concatenated to obtain the full linear convolution result. Each subarray contains $N_1 \cdot N_2 \cdot \ldots \cdot N_M$ values and the $M$-dimension linear convolution has $2N_1 \cdot 2N_2 \cdot \ldots \cdot 2N_M$ values.

If the convolved signals are real, the real and imaginary parts of the weighted circular convolutions are considered in the linear system, and only $2^{M-1}$ convolutions are required to solve the system, reducing the computational costs by about 50 %.

For example, the block diagram illustrated in Fig. 2 presents the operations involved in the case of 3-D real signals.

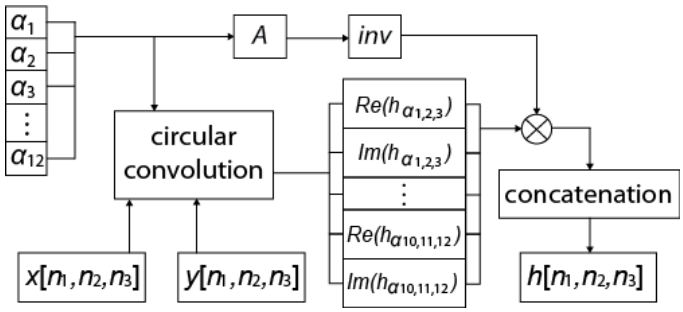


*Fig. C.2. Block diagram of the operations required to calculate 3-D linear convolution when x, y* $\in R$

Here, the linear system requires $2^{3-1} = 4$ weighted circular convolutions to be solved. Consequently, 12 modulation coefficients $\alpha$ are selected (i.e. one coefficient for each dimension in

each convolution). Selecting the values for the modulation coefficients is an important consideration and is discussed in Section III.

Exponential windows are applied to signals $x$ and $y$ and the four weighted circular convolutions are calculated. By taking the real and imaginary parts of the weighted convolution, a system of eight equations is obtained:

$$\begin{bmatrix} Re\{h_{\alpha_{1,2,3}}\} \\ Im\{h_{\alpha_{1,2,3}}\} \\ Re\{h_{\alpha_{4,5,6}}\} \\ Im\{h_{\alpha_{4,5,6}}\} \\ Re\{h_{\alpha_{7,8,9}}\} \\ Im\{h_{\alpha_{7,8,9}}\} \\ Re\{h_{\alpha_{10,11,12}}\} \\ Im\{h_{\alpha_{10,11,12}}\} \end{bmatrix} = [A] \cdot \begin{bmatrix} h_{000} \\ h_{100} \\ h_{010} \\ h_{001} \\ h_{110} \\ h_{101} \\ h_{011} \\ h_{111} \end{bmatrix}, \tag{C.8}$$

The 8 by 8 matrix $A$ contains the real and imaginary components of the exponential modulations with respect to (C.6). By inversing equation (C.8), the linear convolution subarrays can be found. Finally, the subarrays $h_{pqr}$ are concatenated with respect to the notation in (C.7).

The composition of matrix $A$ for the 3-D case is given in appendix A and the 2-D case is described in [96].

## C.5 Choice of modulation coefficients

Selecting proper modulation coefficients is an important step in the application of the proposed method. First, the coefficients must be chosen in order for matrix $A$ to be nonsingular; otherwise, the system cannot be solved. However, the real advantage of this method comes when it is used for multi-dimensional convolutions. By using identical coefficients over the same dimensions in different circular convolutions, the number of FFT calculations can be reduced significantly, while preserving the nonsingularity of matrix $A$. This allows considerable improvement in computational cost.

## Real signals

For example, in the case of 3-D real convolved signals, Table 1 presents values for modulation coefficients that fulfill these requirements. Here, a total of four weighted circular convolutions are required, and each of the 3-D convolutions necessitates three modulation coefficients.

*Tableau C.1: Modulation coefficients for real 3-D signals*

| | $n_1$ | $n_2$ | $n_3$ |
|---|---|---|---|
| $\alpha_{1,2,3}$ | $\pi/2$ | 0 | $\pi$ |
| $\alpha_{4,5,6}$ | $\pi/2$ | 0 | 0 |
| $\alpha_{7,8,9}$ | $\pi/2$ | $\pi$ | 0 |
| $\alpha_{10,11,12}$ | $\pi/2$ | $\pi$ | $\pi$ |

The four triplets are presented; each column corresponds to one of the three dimensions and each row expresses the coefficients for one of the four 3-D circular convolutions. Since the value of $\pi/2$ is used for all coefficients applied to the first dimension, the 1-D FFT calculations involved only have to be computed once for the four circular convolutions. In the second dimension, half of the calculations can be spared in the same manner, since only two 1-D FFT calculations are required over the second dimension.

In fact, in the case of real signals, obtaining linear convolution with the GDFT requires $2^{M-1}$ circular convolutions. In order to express the signals in the *k*-space, these circular convolutions require a total of $M \cdot 2^{M-1}$ one-dimensional generalized Fourier transforms for each of the two convolved signals. Repeating the coefficients as suggested above reduces the number of direct 1D-GDFT from $M \cdot 2^{M-1}$ to only $\sum_{i=0}^{M-1} 2^i = 2^M - 1$.

Fig. 3 illustrates this idea for $M = 3$. By applying the FFT one dimension at the time, the twelve 1-D FFTs applied to $x[n_1, n_2, n_3]$ in the calculation of four 3-D circular convolutions can be reduced to seven because of the repetition of some coefficients.

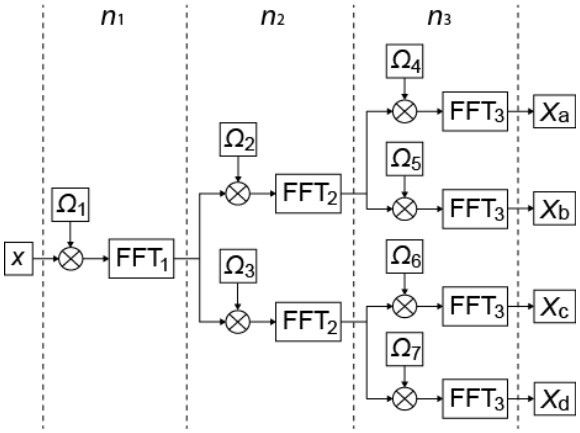


*Figure C.3 : Schematic for the application of the 3D-GDFT on array x with the proposed method*

The four $X$ variables correspond to the data expressed in the 3D-GDFT space and $\Omega_j$ is the exponential window associated to modulation coefficient $\alpha_j$. The coefficients proposed in Tableau C. were also selected in order to avoid calculating a large number of multiplications with exponential window Ω by maximizing the number of null coefficients.

The non-singular matrix $A$ associated with the coefficients shown in Table 1 is presented in (C.9):

$$A = \begin{bmatrix} 1 & 0 & 1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & -1 & 0 & -1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & -1 & 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 & 0 & -1 \\ 1 & 0 & -1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 & -1 & 0 & 1 \end{bmatrix}. \tag{C.9}$$

The condition number associated with A is unitary, and its inverse is:

$$A^{-1} = \frac{1}{4} \cdot \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & -1 & 0 & -1 & 0 \\ -1 & 0 & 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 & 0 & -1 \\ 0 & -1 & 0 & 1 & 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & -1 & 0 & 1 \end{bmatrix}. \tag{C.10}$$

Similar results are obtained for the 2-*D* case with $\alpha_{1,2} = \{\pi/2, \pi\}$ and $\alpha_{3,4} = \{\pi/2, 0\}$, with associated matrix $A$ of unitary condition number.

## Complex signals

In the case of complex convolved signals, $M \cdot 2^M$ 1D-GDFT are required to express each signal in the *k*-space. Consequently, in 3-D, 24 coefficients are required for the calculation of eight circular convolutions. Table 2 proposes a set of values that uses similar duplication scheme to minimize the computational costs.

*Tableau C.2: Modulation coefficients for complex signals*

| | $n_1$ | $n_2$ | $n_3$ |
|---|---|---|---|
| $\alpha_{1,2,3}$ | 0 | 0 | 0 |
| $\alpha_{4,5,6}$ | 0 | 0 | $\pi$ |
| $\alpha_{7,8,9}$ | 0 | $\pi$ | 0 |
| $\alpha_{10,11,12}$ | 0 | $\pi$ | $\pi$ |
| $\alpha_{13,14,15}$ | $\pi$ | 0 | 0 |
| $\alpha_{16,17,18}$ | $\pi$ | 0 | $\pi$ |
| $\alpha_{19,20,21}$ | $\pi$ | $\pi$ | 0 |
| $\alpha_{22,23,24}$ | $\pi$ | $\pi$ | $\pi$ |

Reusing coefficients allows reducing the number of 1*D*-GDFT to $\sum_{i=1}^{M} 2^i = 2^{M+1} - 2$. In 3-D, duplicating the coefficients decreases the number of FFTs required from 24 to 14. The associated matrix $A$ has a unitary condition number. Matrix $A$ and its inverse are presented in (C.11) and (C.12):

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \end{bmatrix} \tag{C.11}$$

$$A^{-1} = \frac{1}{8} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix} \quad \text{(C.12)}$$

For the 2-*D* case, the following coefficients produce a matrix $A$ with a unitary condition number; $\alpha_{1,2} = \{0,0\}$, $\alpha_{3,4} = \{0, \pi\}$, $\alpha_{5,6} = \{\pi, 0\}$, $\alpha_{7,8} = \{\pi, \pi\}$.

## C.6 Analysis of floating point operations

The goal of this section is to evaluate the computational improvement of the multi-dimensional linear convolution due to the duplication of the exponential coefficients presented in Section III. Our approach consists in evaluating the number of floating point operations required by the algorithm, and to compare it to that of the zero-padding and standard GDFT methods. This analysis covers real and complex input signals with an arbitrary number of dimensions $M$. To simplify the analysis, it is supposed that the number of values $N$ is the same for each dimension. For other cases, we suppose the corresponding number of floating point operations is bounded by that of the presented cases, since the computational complexity increases monotonically with $N$.

All studied methods are composed of two main operations: complex multiplication and FFT. The assumptions regarding the computational complexity associated to these operations are described. However, because it is difficult to evaluate the performance of state-of-the-art FFT algorithms, the theoretical analysis presented in this paper is not absolute, and only serves as a guideline for the evaluation of the complexity gains. The results developed here are verified experimentally in the next section by comparing the duration of the computations.

**Fast Fourier transform**

State-of-the-art FFT implementations are optimized for a large variety of cases. For instance, the complexity gain can change with the type of processor used or with the structure of the data. According to Franchetti et al. [88], the fastest FFT libraries on multicore processors require

$4Nlog_2(N)$ to $5Nlog_2(N)$ floating-point operations for complex 1-D signals. A conservative $5Nlog_2(N)$ is considered in the present paper.

In *M*-dimensions, the FFT with respect to a single dimension is applied to all $N^{M-1}$ vectors, resulting in a complexity of $5N^M log_2(N)$ operations. To obtain the *M*-dimensional FFT, this procedure is performed *M* times, once for each dimension. The complexity of the *M*-dimensional FFT (and inverse FFT) is thus evaluated to:

$$C_{FFT}^{\mathbb{C}} = 5MN^M log_2(N). \quad \text{(C.13)}$$

Finally, when applying the FFT to real signals, by exploiting Hermitian symmetry, the calculation times are reduced by half [83]. As an approximation, we thus consider a reduction of half of the floating point arithmetic operations for real signals. The complexity for the *M*-dimensional FFT over real signal becomes:

$$C_{FFT}^{\mathbb{R}} = \frac{5}{2} MN^M log_2(N). \quad \text{(C.14)}$$

This important factor was not considered in previous analyses for GDFT linear convolution efficiency calculation, even though the computational improvement obtained relied on the convolution of real signals [96, 83, 21].

**Complex multiplication**

Complex multiplication requires $6N^M$ floating point operations when both signals are complex, and $3N^M$ floating point operations when one signal is complex and the other is real.

## Complexity for the zero-padding method

The zero-padding method consists in doubling the array in each dimension with zero values before applying circular convolution. In the circular convolution process, two direct FFTs are applied, followed by one inverse FFT. For real convolved signals, the two direct FFTs are applied to real inputs and are evaluated to half the standard operation costs. We assume that the inverse *M*-dimensional FFT is always applied to complex inputs. The two convolved signals are also

multiplied in the *k*-space. Tableau C.3 summarizes the floating point operations required for both real and complex input signals.

*Tableau C.3: Description of the floating point operations required for the zero-padding method*

| | *Operations* | *Number of floating point* |
|---|---|---|
| *Real* | Direct FFTs | $5M(2N)^M \log_2(2N)$ |
| | k-space multiplication | $6(2N)^M$ |
| | Inverse FFT | $5M(2N)^M \log_2(2N)$ |
| *Complex* | Direct FFTs | $10M(2N)^M \log_2(2N)$ |
| | k-space multiplication | $6(2N)^M$ |
| | Inverse FFT | $5M(2N)^M \log_2(2N)$ |

The resulting complexity for this method is thus:

$$C_{zpad}^{\mathbb{R}} = 2^M N^M (10M(log_2(N) + 1) + 6), \tag{C.15}$$

and

$$C_{zpad}^{\mathbb{C}} = 2^M N^M (15M(log_2(N) + 1) + 6), \tag{C.16}$$

for real and complex input signals respectively.

## Complexity for the standard GDFT method

The number of floating point operations for the standard GDFT method without any duplication of the modulation coefficients is evaluated. Compared to the zero-padding method, the GDFT method has the following computing advantages:

i. The processed signals have half the number of values in each dimension;

ii. The number of GDFT convolutions is reduced by half when the input signals are real.

Three types of operations are considered:

**Direct and inverse FFT**

As discussed in Section III, $2^{M-1}$ circular convolutions are required for real signals, and twice as many for complex signals. Each circular convolution requires tree $M$-dimensional FFT times (two direct FFTs and one inverse).

**Direct and inverse exponential windows**

Each circular convolution requires three multiplications with exponential windows (one for each direct or inverse MD-GDFT applied), resulting in $4 \cdot 2^{M-1} 3N^M$ operations for real signals, and $3 \cdot 2^M 6N^M$ for complex signals. The direct modulations applied to real signals require only $3N^M$ operations.

**Multiplication in the k-space**

Multiplication of the signals in the *k*-space is performed once for every circular convolution.

The operations are summarized in Tableau C.4.

*Tableau C.4: Description of the floating point operations required for the standard GDFT method*

| | *Operations* | *Number of floating point operations* |
|---|---|---|
| *Real* | Direct modulations | $6 \cdot 2^{M-1} N^M$ |
| | Direct FFTs | $10M \cdot 2^{M-1} N^M \log_2(N)$ |
| | k-space multiplication | $6 \cdot 2^{M-1} N^M$ |
| | Inverse FFT | $5M \cdot 2^{M-1} N^M \log_2(N)$ |
| | Inverse modulations | $6 \cdot 2^{M-1} N^M$ |
| *Complex* | Direct modulations | $12 \cdot 2^M N^M$ |
| | Direct FFTs | $10M \cdot 2^M N^M \log_2(N)$ |
| | k-space multiplication | $6 \cdot 2^M N^M$ |
| | Inverse FFT | $5M \cdot 2^M N^M \log_2(N)$ |
| | Inverse modulations | $6 \cdot 2^M N^M$ |

Therefore, the sum of the floating point operations for real and complex signals is:

$$C_{GDFT}^{\mathbb{R}} = 2^{M-1}N^M(15Mlog_2(N) + 18), \quad \text{(C.17)}$$

$$C_{GDFT}^{\mathbb{C}} = 2^M N^M(15Mlog_2(N) + 24). \quad \text{(C.18)}$$

## Complexity for the GDFT method with duplications

The same types of operations are considered when applying duplications of the modulation coefficients. Compared to the standard GDFT method, the main computational advantage of the duplication method is the reduction of the number of direct 1-D FFT by half minus one. Again, three types of operations are considered:

**Direct and inverse FFT**

As discussed in Section IV, for real inputs, $2^M - 1$ one-dimensional FFT are required in order to express each signal in the *k*-space. This process is performed twice, once for each signal. The duplication scheme cannot be applied for inverse transforms, which require $M \cdot 2^{M-1}$ FFT. For complex signals, twice as many operations are required.

**Direct and inverse exponential windows**

Each FFT application requires a multiplying with an exponential window. Here, it is assumed conservatively that no modulation coefficient is null. Direct modulation of real signals for the first dimension requires half of the operations.

**Multiplication in the k-space**

Multiplying signals in the *k*-space is performed once for every circular convolution.

The operations are summarized in Tableau C.5.

*Tableau C.5: Description of the floating point operations required for the method with duplications*

| | *Operations* | *Number of floating point operations evaluated* |
|---|---|---|
| *Real* | Direct modulations | $6 \cdot (2^{M+1} - 3)N^M$ |
| | Direct FFTs | $10(2^M - 1) \cdot N^M \log_2(N)$ |
| | k-space multiplication | $6 \cdot 2^{M-1}N^M$ |
| | Inverse FFT | $5M \cdot 2^{M-1}N^M \log_2(N)$ |
| | Inverse modulations | $6 \cdot 2^{M-1}N^M$ |
| *Complex* | Direct modulations | $12 \cdot (2^{M+1} - 2)N^M$ |
| | Direct FFTs | $10(2^{M+1} - 2) \cdot N^M \log_2(N)$ |
| | k-space multiplication | $6 \cdot 2^M N^M$ |
| | Inverse FFT | $5M \cdot 2^M N^M \log_2(N)$ |
| | Inverse modulations | $6 \cdot 2^M N^M$ |

Hence, the total number of operations is estimated to:

$$C_d^{\mathbb{R}} = (2^M - 1)10N^M log_2(N) + 2^{M-1}N^M(5Mlog_2(N) + 12) + 6(2^{M-1} - 3)N^M, \quad \text{(C.19)}$$

and for complex signals:

$$C_d^{\mathbb{C}} = (2^{M+1} - 2)(10N^M log_2(N) + 12) + 2^M N^M(5Mlog_2(N) + 12). \quad \text{(C.20)}$$

## Comparison and analysis

The resulting computational complexities as a function of *N* are illustrated for the cases of two-dimensional and three-dimensional real and complex signals. Comparison is also made with and without the duplicating the modulation coefficients. The results are presented in Figure 4 for the 2-D case and in Figure 5 for the 3-D case.

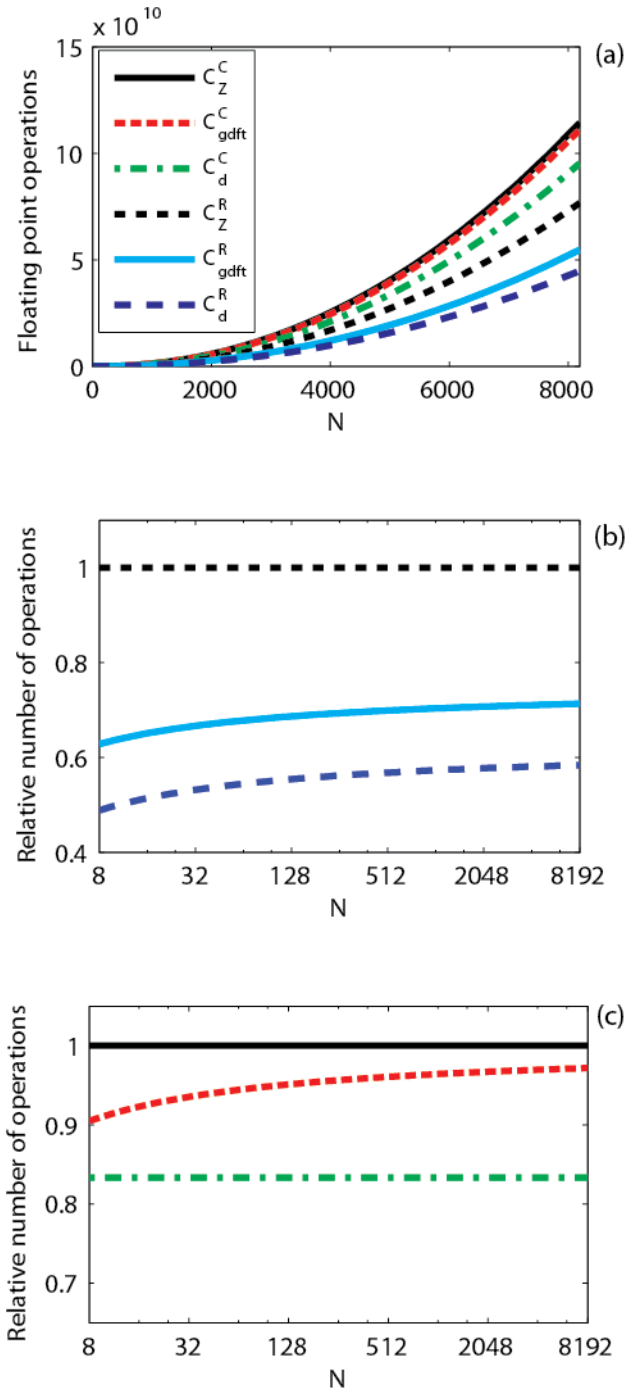


*Figure C.4 : 2-D comparison of computational complexity of the proposed method with real and complex convolved signals, along with standard GDFT and zero-padding method; a) Absolute complexity. b) Complexity relative to the zero-padding method for real signals. c) Complexity relative to the zero-padding method for complex signals. The legend is the same for all frames*

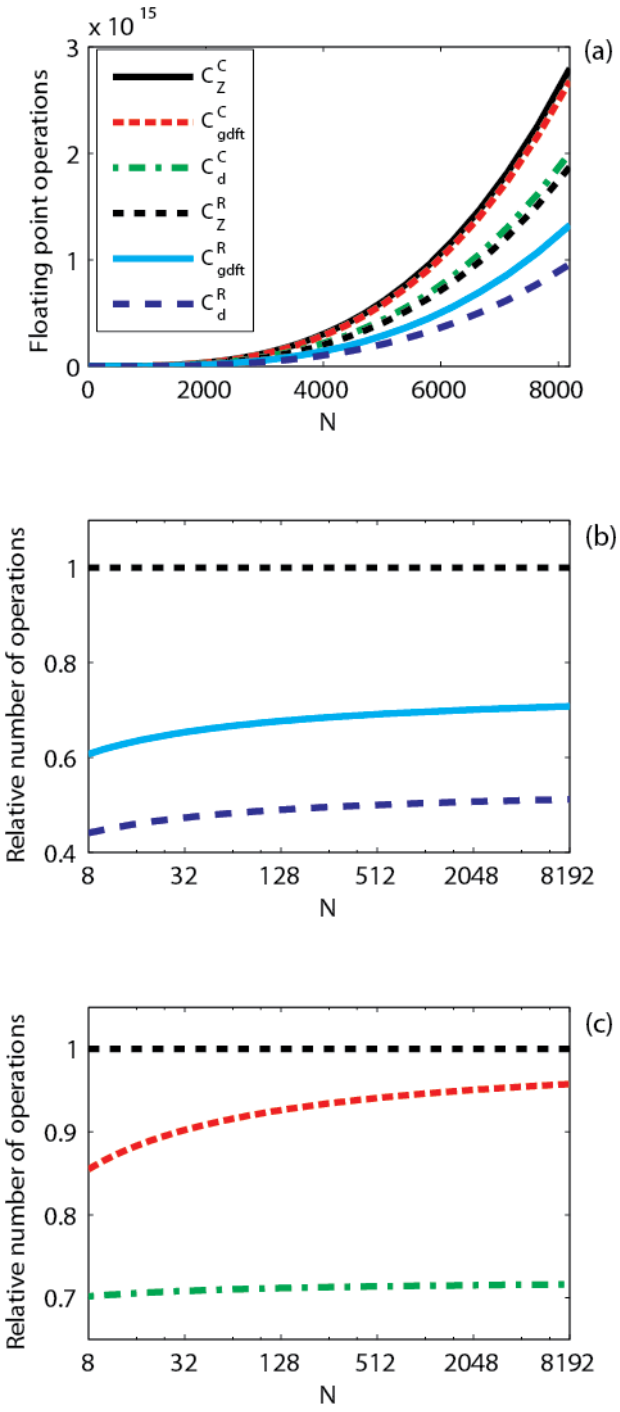


*Figure C.5 : 3-D comparison of computational complexity of the proposed method with real and complex convolved signals, along with standard GDFT and zero-padding method; a) Absolute complexity. b) Complexity relative to the zero-padding method for real signals. c) Complexity relative to the zero-padding method for complex signals. The legend is the same for all frames*

Figure C.4 b, Figure C.4 c, Figure C.5 b and Figure C.5 c show the complexity normalized with respect to the zero-padding method. For 2-D real signals, the duplication technique reduces the complexity to about half of that of the zero-padding method, which corresponds to an improvement of about 12 % under that of the standard GDFT method. For 2-D complex signals, the improvement is between 7 % to 14 % depending on the number of values $N$. Compared to the standard GDFT method, the complexity reduction is considerable when convolving complex signals, especially for larger values of $N$. In 3-D, the reduction is around 17 % to 29 % under that of the standard GDFT method for real convolved signals, and from 15 % to 24 % for complex signals. The reduction in computational complexity increases with the number of dimensions, as supported by (C.15) to (C.20). In all studied cases, the application of the proposed algorithm significantly reduces the complexity of the calculations, even when compared to the standard GDFT algorithm.

The relative complexity ratio of the standard GDFT method reaches about 70 % in the case of real signals. This result does not corroborate exactly the analysis of Radakrishnan and Jenkins for the 2-D case which predicted a reduction of 50 % with respect to the zero-padding method [9], since we considered the reduced cost for applying the FFT to real signals. For complex signals, the complexity without duplication of the coefficients converges to that of the zero-padding method. However, with the duplication technique, the complexity decreases to around 85 % and 70 % compared to the zero-padding method, in 2-D and 3-D respectively.

In this analysis, parallelism is not taken into account. However, since each circular convolution can be calculated independently, processing of the convolutions simultaneously in different channels can certainly be an asset. Additionally, the analysis does not consider cases where the modulation parameter is null: this would imply that the computational cost of the multiplication with the unitary window is avoided.

## C.7 Computational validation

In order to support these predictions, the linear convolution of randomly generated signals in single-precision floating-point format is computed with the standard GDFT and duplication methods. It is convenient to compare these two methods experimentally, since the type of operations and signals involved are similar, and the processing should be handled the same way. The calculation times $t_{GDFT}$ and $t_d$ are obtained for both methods, and the ratio $t_d/t_{GDFT}$ is

calculated, along with the complexity ratio $C_d/C_{GDFT}$, in order to compare both methods. The results are presented with respect to the number of values $N$ in Figure C.6 for the 2-D and 3-D cases with real and complex values; it is supposed that $N$ is the same for each dimension.

In order to preserve a low variance of the results, the calculation times are obtained from the average of enough computations for the total calculation time to be at least over one second. The calculations are performed on Python using the FFT function included in the *Numpy* module [97]. An i7-2600 Intel 3.40 GHz processor was used for the calculations.

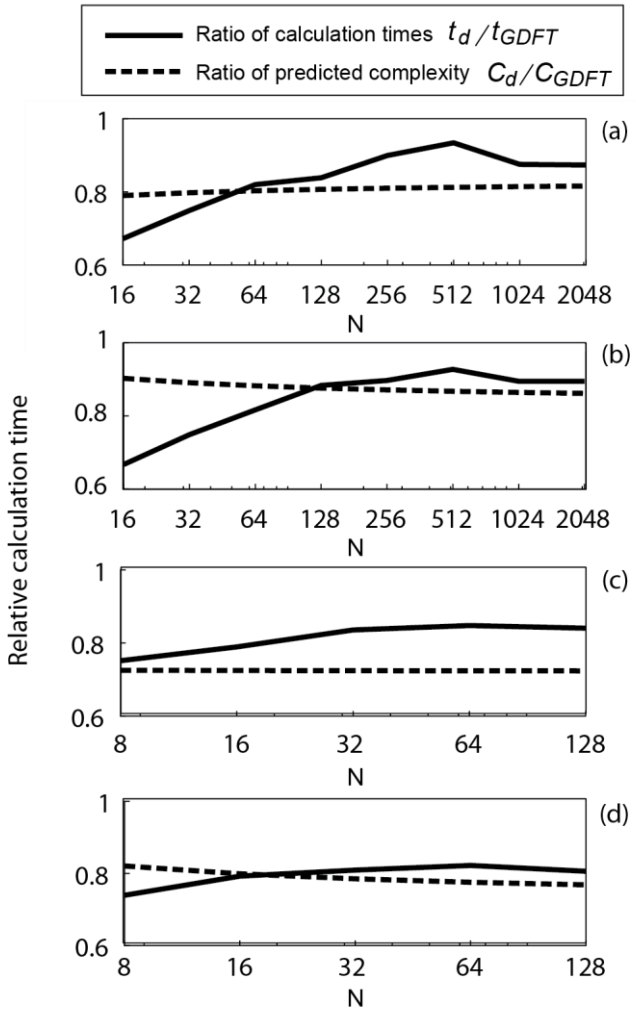


*Figure C.6 : Relative efficiency of the duplication method with respect to the standard GDFT method for (a). 2-D real signal (b) 2-D complex signal (c) 3-D real signal (d) 3-D complex signal*

In Fig. 6, the calculation results are generally in good agreement with the theoretical predictions. In all cases studied, the benefits of the duplication method over the standard GDFT is a bit overestimated by the complexity analysis when the signals have a higher number of data. Still, the floating point operations analysis gives a good estimate of the calculation times: the average difference remains under 9 % for all presented cases.

The results show that the duplication algorithm reduces the calculation times by about 15 to 20 % relatively to that of the GDFT algorithm without duplication. This confirms that the proposed algorithm is relevant to the computation of multi-dimensional linear convolutions.

## C.8 Conclusion

An algorithm based on the concatenation of weighted circular convolution is used to efficiently compute multi-dimensional linear convolution. It is shown that by appropriately choosing the modulation coefficients associated with multi-dimensional weighted circular convolution, it is possible to considerably decrease the number of direct Fourier transforms required. Sets of modulation coefficients are proposed for both real and complex convolved signals in the two-dimensional and three-dimensional cases. This reduces the total computational complexity in the order of 10 % to 25 % of that of the standard zero-padding method. This reduction is verified with experimental computation times.

Prospective work could include an analysis of how the proposed method benefits from parallel computation.

## C.9 Appendix A

The eight-by-eight matrix $A$ for 3-D linear convolution of real signals is:

$$A = \begin{bmatrix} 1 & R_1 & R_2 & R_3 & (R_1R_2 - I_1I_2) & (R_1R_3 - I_1I_3) & (R_2R_3 - I_2I_3) \\ 0 & I_1 & I_2 & I_3 & (R_1I_2 + I_1R_2) & (R_1I_3 + I_1R_3) & (R_2I_3 + I_2R_3) \\ 1 & R_4 & R_5 & R_6 & (R_4R_5 - I_4I_5) & (R_4R_6 - I_4I_6) & (R_5R_6 - I_5I_6) \\ 0 & I_4 & I_5 & I_6 & (R_4I_5 + I_4R_5) & (R_4I_6 + I_4R_6) & (R_5I_6 + I_5R_6) \\ 1 & R_7 & R_8 & R_9 & (R_7R_8 - I_7I_8) & (R_7R_9 - I_7I_9) & (R_8R_9 - I_8I_9) \\ 0 & I_7 & I_8 & I_9 & (R_7I_8 + I_7R_8) & (R_7I_9 + I_7R_9) & (R_8I_9 + I_8R_9) \\ 1 & R_{10} & R_{11} & R_{12} & (R_{10}R_{11} - I_{10}I_{11}) & (R_{10}R_{12} - I_{10}I_{12}) & (R_{11}R_{12} - I_{11}I_{12}) \\ 0 & I_{10} & I_{11} & I_{12} & (R_{10}I_{11} + I_{10}R_{11}) & (R_{10}I_{12} + I_{10}R_{12}) & (R_{11}I_{12} + I_{11}R_{12}) \end{bmatrix}$$

$$\begin{bmatrix} (R_1R_2R_3 - I_1I_2R_3 - R_1I_2I_3 - I_1R_2I_3) \\ (R_1I_2R_3 + I_1R_2R_3 + R_1R_2I_3 - I_1I_2I_3) \\ (R_4R_5R_6 - I_4I_5R_6 - R_4I_5I_6 - I_4R_5I_6) \\ (R_4I_5R_6 + I_4R_5R_6 + R_4R_5I_6 - I_4I_5I_6) \\ (R_7R_8R_9 - I_7I_8R_9 - R_7I_8I_9 - I_7R_8I_9) \\ (R_7I_8R_9 + I_7R_8R_9 + R_7R_8I_9 - I_7I_8I_9) \\ (R_{10}R_{11}R_{12} - I_{10}I_{11}R_{12} - R_{10}I_{11}I_{12} - I_{10}R_{11}I_{12}) \\ (R_{10}I_{11}R_{12} + I_{10}R_{11}R_{12} + R_{10}R_{11}I_{12} - I_{10}I_{11}I_{12}) \end{bmatrix} \quad \text{(C.21)}$$

$R_m$ and $I_m$ are respectively the real and imaginary components of the modulation exponential $e^{i\alpha_m}$, with $m \in \{1, \dots, 12\}$.

In the case of complex convolved signals, the matrix $A$ is:

$$A = \begin{bmatrix} 1 & e^{i\alpha_1} & e^{i\alpha_2} & e^{i\alpha_3} & e^{i(\alpha_1+\alpha_2)} & e^{i(\alpha_1+\alpha_3)} & e^{i(\alpha_2+\alpha_3)} & e^{i(\alpha_1+\alpha_2+\alpha_3)} \\ 1 & e^{i\alpha_4} & e^{i\alpha_5} & e^{i\alpha_6} & e^{i(\alpha_4+\alpha_5)} & e^{i(\alpha_4+\alpha_6)} & e^{i(\alpha_5+\alpha_6)} & e^{i(\alpha_4+\alpha_5+\alpha_6)} \\ 1 & e^{i\alpha_7} & e^{i\alpha_8} & e^{i\alpha_9} & e^{i(\alpha_7+\alpha_8)} & e^{i(\alpha_7+\alpha_9)} & e^{i(\alpha_8+\alpha_9)} & e^{i(\alpha_7+\alpha_8+\alpha_9)} \\ 1 & e^{i\alpha_{10}} & e^{i\alpha_{11}} & e^{i\alpha_{12}} & e^{i(\alpha_{10}+\alpha_{11})} & e^{i(\alpha_{10}+\alpha_{12})} & e^{i(\alpha_{11}+\alpha_{12})} & e^{i(\alpha_{10}+\alpha_{11}+\alpha_{12})} \\ 1 & e^{i\alpha_{13}} & e^{i\alpha_{14}} & e^{i\alpha_{15}} & e^{i(\alpha_{13}+\alpha_{14})} & e^{i(\alpha_{13}+\alpha_{15})} & e^{i(\alpha_{14}+\alpha_{15})} & e^{i(\alpha_{13}+\alpha_{14}+\alpha_{15})} \\ 1 & e^{i\alpha_{16}} & e^{i\alpha_{17}} & e^{i\alpha_{18}} & e^{i(\alpha_{16}+\alpha_{17})} & e^{i(\alpha_{16}+\alpha_{18})} & e^{i(\alpha_{17}+\alpha_{18})} & e^{i(\alpha_{16}+\alpha_{17}+\alpha_{18})} \\ 1 & e^{i\alpha_{19}} & e^{i\alpha_{20}} & e^{i\alpha_{21}} & e^{i(\alpha_{19}+\alpha_{20})} & e^{i(\alpha_{19}+\alpha_{21})} & e^{i(\alpha_{20}+\alpha_{21})} & e^{i(\alpha_{19}+\alpha_{20}+\alpha_{21})} \\ 1 & e^{i\alpha_{22}} & e^{i\alpha_{23}} & e^{i\alpha_{24}} & e^{i(\alpha_{22}+\alpha_{23})} & e^{i(\alpha_{22}+\alpha_{24})} & e^{i(\alpha_{23}+\alpha_{24})} & e^{i(\alpha_{22}+\alpha_{23}+\alpha_{24})} \end{bmatrix}, \quad \text{(C.22)}$$

with $m \in \{1, \dots, 24\}$.